

How faculty define quality, prestige, and impact in research

Esteban Morales^{*1}, Erin McKiernan², Meredith T. Niles³, Lesley Schimanski⁴, Juan Pablo Alperin^{*4}

1. University of British Columbia, Canada.
2. Universidad Nacional Autónoma de México
3. University of Vermont
4. Simon Fraser University

*Corresponding authors

Abstract

Despite the calls for change, there is significant consensus that when it comes to evaluating publications, review, promotion, and tenure processes should aim to reward research that is of high "quality," has an "impact," and is published in "prestigious" journals. Nevertheless, such terms are highly subjective and present challenges to ascertain precisely what such research looks like. Accordingly, this article responds to the question: how do faculty from universities in the United States and Canada define the terms quality, prestige, and impact? We address this question by surveying 338 faculty members from 55 different institutions. This study's findings highlight that, despite their highly varied definitions, faculty often describe these terms in overlapping ways. Additionally, results shown that marked variance in definitions across faculty does not correspond to demographic characteristics. This study's results highlight the need to more clearly implement evaluation regimes that do not rely on ill-defined concepts.

Financial Disclosure

Funding for this project was provided to JPA, MTN, ECM, and LAS from the OpenSociety Foundations (OR2017-39637). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Related Materials

Other publications related to this project, including a series of infographics summarizing findings, can be found at: <https://www.scholcommlab.ca/research/rpt-project/>

Survey responses can be found at the following publication:

Niles, Meredith T.; Schimanski, Lesley A.; McKiernan, Erin C.; Alperin, Juan Pablo, 2020, "Data for: Why we publish where we do", <https://doi.org/10.7910/DVN/MRLHNO>, *Harvard Dataverse*, V1

Data regarding RPT documents can be found at the following data publication:

Alperin, Juan Pablo; Muñoz Nieves, Carol; Schimanski, Lesley; McKiernan, Erin C.; Niles, Meredith T., 2018, "Terms and Concepts found in Tenure and Promotion Guidelines from the US and Canada", <https://doi.org/10.7910/DVN/VY4TJE>, *Harvard Dataverse*, V3, UNF:6:PQC7QoilolhDrokzDPxyQ== [fileUNF]

1

Introduction

2 Although faculty work involves a wide range of activities and priorities (Fox, 2012; Gruzd
3 et al., 2011; Miller & Seldin, 2014), evidence confirms that faculty, even those at teaching and
4 public institutions, believe that research is the most highly valued aspect of their work (Acker &
5 Webber, 2016; Bergeron et al., 2014; Harley et al., 2010). It is therefore unsurprising that
6 despite significant efforts to broaden how university faculty are assessed, research continues to
7 be treated as the main component of a professor's job (Chen, 2015; Gordon, 2008; Harley et al.,
8 2010) and, accordingly, research assessment features prominently in how faculty are evaluated
9 for career advancement. Yet, despite its perceived importance and critical role, there is still
10 debate about how to evaluate research outputs.

11 Research assessment, which is codified for faculty career progression in review,
12 promotion, and tenure (RPT) processes, is itself a controversial topic that has been subject to
13 much debate over the years (c.f., Schimanski & Alperin, 2018). Critics have argued that current
14 assessment practices fail to recognize new and evolving forms of scholarship, such as datasets
15 and research software (Genshaft et al., 2016; Howard, 2013; Piwowar, 2013; Sanberg et al.,
16 2014), fail to encourage reproducible science (Open Science Collaboration, 2015), or worse,
17 that current approaches encourage unethical practices, such as gift authorship, p-value hacking,
18 or result manipulation (Chapman et al., 2019; Edwards & Roy, 2017). These discussions point
19 to different perspectives regarding what constitutes research that is worthy of being celebrated
20 and rewarded.

21 Despite the calls for change and the differences of opinions, there is significant
22 consensus that when it comes to evaluating publications, the RPT process should aim to reward
23 research that is of high "quality" and has "impact" (Schimanski & Alperin, 2018). However, such
24 terms are highly subjective, and present challenges to ascertain precisely what such research
25 looks like. Furthermore, their subjectivity presents additional challenges for comparing research
26 and individuals, as is regularly done during the RPT process. The use of these subjective
27 concepts, and others like them, may serve primarily as "rhetorical signalling device[s] used to
28 claim value across heterogeneous institutions, researchers, disciplines, and projects rather than
29 a measure of intrinsic and objective worth" (Moore et al., 2017, p. 3).

30 Others have previously noted the lack of clear definitions surrounding many of the terms
31 and concepts used in research assessment (Dean et al., 2016; Hatch, 2019; Moore et al., 2017;

1 van Mil & Henman, 2016). Without definitions, individuals and committees are bound to apply
2 different standards which inevitably leads to inequities in how faculty and research are
3 evaluated. If, as Hatch (2019) suggests, “most assessment guidelines permit sliding standards,”
4 then the definition used in any given situation can easily shift depending on whose work is being
5 assessed in ways that allow (or even encourage) biases to creep into the evaluation process
6 (be they conscious or unconscious). Even if individuals are consistent and unbiased in how they
7 apply their own definitions, there is rarely agreement in assessments between academics, even
8 from those of the same discipline. Moore and colleagues (2017) point to such conflicting
9 assessments as the clearest example of a lack of any agreed upon definition of ‘excellence’—an
10 argument that can easily be extended to other terms and concepts commonly used in research
11 assessment.

12 The known pernicious effects of using ill-defined criteria has resulted in calls to “swap
13 slogans for definitions” (Hatch, 2019) and for greater “conceptual clarity” (Belcher & Palenberg,
14 2018) in research assessment. To aid in this effort, this study asks the question: how do faculty
15 understand the concepts of ‘quality’, ‘impact’, and ‘prestige’ as they pertain to RPT processes?

16 **Previous research**

17 The current study builds on our previous work to study the documents related to RPT
18 processes at institutions in the United States and Canada (Alperin et al., 2019; McKiernan et al.,
19 2019; Niles et al., 2020), as well as similar work by others in this area (Moher et al., 2018; Rice
20 et al., 2020; Snider et al., 2020). In one of our previous studies (Alperin et al., 2019), we
21 reported that nearly 60% of institutions overall and nearly 80% of research-intensive universities
22 mentioned ‘impact’ in their RPT documents. Using the same data and analysis, but previously
23 unreported, we also found that 73% of institutions in our sample and 89% of research-intensive
24 universities mentioned ‘quality’ in their RPT documents, and that those percentages were 29%
25 and 47% for those that mentioned ‘prestige.’ That is to say, there was a high prevalence of such
26 concepts in academic evaluations, but a closer reading of these instances shows that few of
27 these documents gave a clear definition of what these terms meant or how they were to be
28 measured.

29 Despite the frequent use of these terms in relation to research assessment, it is difficult
30 to know how faculty understand them, especially in relation to how they will assess others, and
31 in how they understand that they will be assessed. A survey by DeSanto and Nichols (2017)

1 found that "...a significant number of faculty [are] unsure of their department's RPT expectations
2 for demonstrating scholarly impact" (pg. 156). Results from that same study show there is
3 substantial disagreement among faculty as to how impact should be measured and evaluated,
4 with many pushing for traditional journal-level metrics like the Journal Impact Factor (JIF) and a
5 small percentage favoring new article-level metrics and altmetrics. Similarly, there is
6 disagreement as to how research quality should be measured, with little evidence suggesting
7 that it can be assessed through citation-based metrics like the JIF (Aksnes et al., 2019). In lieu
8 of an objective measure of quality, it has become common to use the perceived prestige of the
9 publication venue (i.e., the journal where an article is published) as a proxy. To confound things
10 further, prestige is itself sometimes associated with the JIF (e.g., Vinyard & Colvin, 2018), even
11 while the association of the JIF with both quality and prestige has been heavily criticized, most
12 notably in The San Francisco Declaration on Research Assessment (Cagan, 2013; DORA,
13 n.d.), the HuMetricsHSS Initiative (n.d.) and in the Leiden Manifesto (Hicks et al., 2015).

14 This interplay between the JIF, quality, impact, and prestige and how it features in
15 research assessment is also evident when faculty make decisions about where to publish their
16 research. A recent faculty survey (Blankstein & Wolff-Eisenberg, 2019) shows that nearly 80%
17 of faculty report the JIF as one of the most important factors influencing their decisions on
18 where to publish, something echoed in our own survey findings (Niles et al., 2020). In some
19 instances, faculty are guided by librarians to use the JIF to determine the prestige of a given
20 journal (Vinyard & Colvin, 2018) but, according to the same Ithaka survey, the majority of faculty
21 make such decisions based on their own perceptions of the quality, prestige, and potential
22 impact of a given journal. As the report states: "less than 20% of respondents reported they
23 receive help...determining where to publish to maximize impact, and assessing impact following
24 publication" (Blankstein & Wolff-Eisenberg, 2019, p. 27).

25 These surveys show that, when it comes to making decisions about where to publish,
26 faculty see an interplay between the notions of impact, quality and prestige and have
27 themselves linked these to metrics like the JIF, although their precise understanding of these
28 terms remains unclear. To some extent, these interconnected concepts have been codified in
29 the RPT guidelines and documents that govern academic careers. As noted above, in previous
30 work we uncovered the high incidence of the terms *quality* and *impact* in these documents
31 (Alperin et al., 2019) and, in another study, we uncovered that the JIF and related terms are
32 found in the RPT documents of 40% of R-type institutions, and the overwhelming majority of
33 those mentions support their use (McKiernan et al., 2019). Similarly, Rice et al. (2020) found

1 mentions of the JIF in nearly 30% of the RPT guidelines from several countries, and also found
2 support for the measure's use. Moreover, we found that although it is not always stated what the
3 JIF is intended to measure, 63% of institutions that mentioned the JIF in their documents had at
4 least one instance of associating the metric with quality, 40% had at least one mention
5 associating it with impact, and 20% with prestige (McKiernan et al., 2019). These results are in
6 stark contrast to a number of studies showing that JIF has little or nothing to do with research
7 quality (Brembs, 2018; Fraley & Vazire, 2014; Munafò et al., 2009; Szucs & Ioannidis, 2017).

8 The complexity of these terms, intertwined with their application in faculty behavior and
9 promotion decisions, demonstrate a need to further understand how faculty themselves
10 perceive these terms. Their persistent use in both publication decisions and in research
11 assessment indicates their importance and the ambiguities, regardless of their reason for being,
12 suggest that further study is needed. As such, we sought to answer the question: how do faculty
13 from universities in the United States and Canada define the terms quality, prestige, and
14 impact?

15 **Methods**

16 As part of the larger project (Alperin et al., 2019; McKiernan et al., 2019; Niles et al.,
17 2020), we collected documents related to the RPT process from a random sample of
18 universities in the United States and Canada. We describe the methods and considerations
19 used when selecting institutions, the process for identifying the documents, and the approach
20 used to identify the presence of terms in Alperin et al. (2019) and in the methodological note
21 accompanying the public dataset (Alperin et al., 2018). To summarize, we followed a stratified
22 sampling technique to identify a representative set of universities from two lists: the 2015 edition
23 of the Carnegie Classification of Institutions (n.d.) and the 2016 edition of the Maclean's
24 Rankings (n.d.). Both lists, using slightly different definitions and nomenclature, allowed us to
25 draw from three types of institutions: those that award many doctoral degrees and are research-
26 intensive (R-type), those that award many master's degrees but few or no doctoral degrees (M-
27 type), and those that award few graduate but many undergraduate (baccalaureate) degrees (B-
28 type), and those that focus on undergraduate programs (i.e., baccalaureate; B-type). We then
29 used a combination of web searches, requests on social media, and targeted emailing to collect
30 documents related to the RPT process.

1 Through this process, we collected a total of 859 documents, including faculty
2 handbooks, collective agreements, memos and forms. The documents corresponded to 129
3 institutions, of which 57 were R-type, 39 were M-type, and 33 were B-type institutions. While
4 some documents applied to the institution as a whole (e.g., collective agreements), other
5 documents were from a specific academic unit (i.e., a department or faculty). The final set
6 contained documents from 381 distinct academic units, themselves a part of 60 of 129
7 universities.

8 As described in more detail in McKiernan et al., (2019), we used text queries with QSR
9 International's NVivo 12 qualitative data analysis software to search every document for terms
10 related to the JIF, including "impact factor", "impact score", "impact metric", "impact index",
11 "high-impact journal", "impact of the journal" and "journal('s) impact." We considered that an
12 academic unit mentions the JIF if we were able to find any of the related terms mentioned in at
13 least one of the documents from that unit.

14 In addition, and as described in greater detail in Niles et al. (2020), we took the list of the
15 381 academic units for which we had at least one document and conducted a web search for a
16 page containing a listing of faculty members from each unit. From these, we picked up to five
17 members from the list without considering the individual's characteristics or their position in the
18 list. We were able to identify 1,644 faculty email addresses from 334 of the 381 academic units.
19 Although some units did not have their faculty email addresses listed publicly, and others did not
20 have 5 faculty members listed, we were able to obtain email addresses for at least one unit at
21 each of the 60 universities on our list.

22 We took this list of email addresses and invited faculty to participate in a survey between
23 September and October 2018. We received responses from 338 faculty (22%) from 55 different
24 institutions. Of these, 84 (25%) were faculty at Canadian institutions and the remaining 254
25 (75%) were from the United States; 223 (66%) were from R-Type institutions, 111 (33%) from
26 M-Type institutions, and 4 (1%) from B-Type institutions. Full methodological details and
27 demographic reporting of respondents can be found in Niles et al. (2020).

28 In this paper, we present a detailed analysis of the 336 responses to the question "*In*
29 *your own words, how would you define the following terms, sometimes used to describe*
30 *academic journals?*" The terms included in this question were *high quality*, *prestigious* and *high*
31 *impact*. We analyzed the responses using open-coding and constant comparison (Strauss &

1 Corbin, 1990). To achieve this, we first organized all the responses into segments—sentences,
2 or parts of sentences, that convey a single idea. We then assigned codes to these segments,
3 grouping those that represent the same idea or creating new codes when a new idea appeared.
4 Each response could contain one or multiple segments, each of which could be coded
5 differently, allowing for a single response to have multiple codes. This process continued until
6 we developed a codebook that included the name of the codes, descriptions for each and
7 examples, as shown in *Tables 1, 2 and 3* (below, in the results section).

8 In order to determine the inter-rater reliability of the codebook, two researchers
9 independently coded the same randomly chosen set of 20 responses for each of the three
10 terms, and compared the results using NVivo 12, which resulted in adjustments to the codebook
11 and finally an overall Kappa value of 0.87 (McHugh, 2012). After having a good result in the
12 inter-rater reliability test, all the responses were coded by one of the authors of this study (EM).
13 Finally, the results of the open-coding process were analyzed by running Chi-Square tests upon
14 different variables in Excel, in order to see the variation of the definitions in the light of different
15 categories.

16 **Results**

17 We present the results in three parts: first, we describe the results of the open-coding
18 process for the three terms, as the codes themselves capture the definitions used by
19 respondents; second, we analyze the frequencies of each code in relation to respondent's
20 demographic information; and, finally, given the high incidence of the JIF in faculty definitions of
21 the three terms, we explore the relationship between faculty definitions and the presence of the
22 JIF in the RPT guidelines at the faculty member's institution and academic unit.

23 **Defining quality, prestige and impact**

24 The analysis of the definitions provided for *high quality* resulted in 295 segments. These
25 segments were categorized into five groups: *Impact factor and Metrics, Value, Readership,*
26 *Reputation, and Review Process*. Table 1 provides the description of each group, as well as
27 some examples for each of them.

28 The result of this coding process shows that faculty most commonly define high quality
29 based on the review process, referring to the perceived rigor of the process of evaluating,
30 gatekeeping and editing academic articles for the journal (e.g., "rigorous review and selection of

1 articles to publish” [848]) (Figure 1). Another common view on what determines the quality of an
 2 academic journal is related to the perceived value of the articles published within it, including
 3 how consequential they are for the field, the quality of the writing, and the methodological
 4 standards of the research (e.g., “Good methodology, quality writing” [1624]).

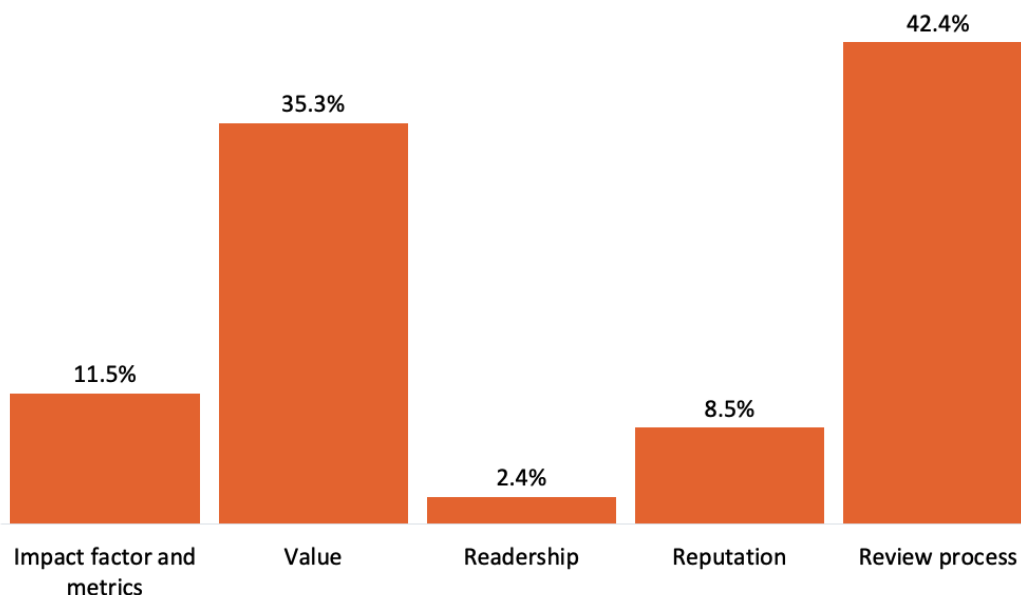
5 **Table 1.** Categories identified in the definitions of High Quality.

6 *Name, description and examples of the categories found in participants’ definitions of High Impact. Numbers in*
 7 *square brackets represent anonymized identification of participants.*

Category	Description of the category	Examples
Impact factor and metrics	Established measurement for the journals by the articles that are published, and the citation generated	<ul style="list-style-type: none"> • High journal ranking [373] • Journal impact factor [500]
Value	Quality and applicability of the articles published in the journal, including its scientific rigor and contribution to the field	<ul style="list-style-type: none"> • Quality of the research being presented [582] • Publishes well-researched, innovative articles [703]
Readership	Focuses on how much the published work is read by academic and non-academic people	<ul style="list-style-type: none"> • Large readership [62] • Is widely read [621]
Reputation	Influence and recognition of all the elements related to the journal, such as the editorial board, the scholars who publish or the journal itself	<ul style="list-style-type: none"> • Well-established, well-known editorial board [568] • name recognition of journal [162]
Review process	Elements related to the process of reviewing the articles that are published, such as editors, feedback and rejection rate.	<ul style="list-style-type: none"> • Peer-reviewed publication [1610] • Rigorous, reviewed by top reviewers in the field [600]

8

1



2

3 **Figure 1. Definitions of High Quality.**

4 *Percentage of responses that contained at least one segment in each of the categories identified in participants'*
5 *definitions of High Quality.*

6 The analysis of the definitions provided for *prestige* resulted in 262 segments. These
7 segments were coded under six categories: *Impact Factor and Metrics*, *Quality and Relevance*,
8 *Readership*, *Relation to Associations*, *Reputation*, and *Review Process*. Table 2 provides
9 definitions for each of the categories, as well as some examples for each of them.

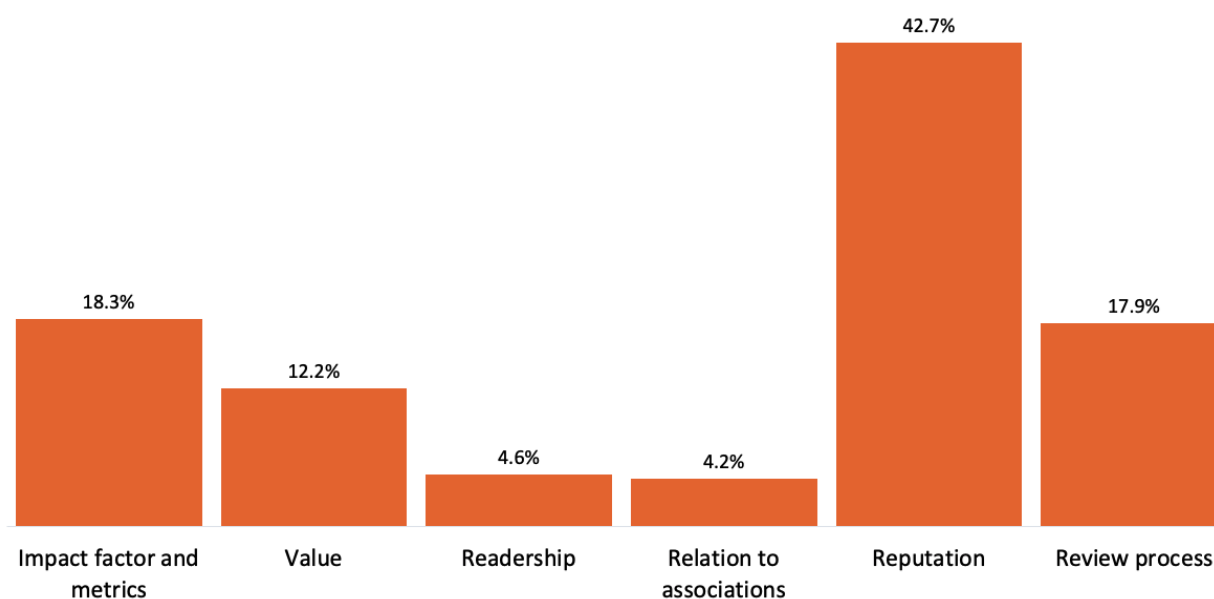
10 The result of this coding process shows that the prestige of academic journals is, in a
11 somewhat circular fashion, most commonly defined by their reputation, which is related to the
12 name recognition of the journal, the people who publish in it or the people in charge of the
13 review process (Figure 2). This was exemplified by definitions like “well-regarded by others in a
14 field” [719], “the journal has a known name in my field of study” [1565] or “well regarded with
15 global recognition” [827]. *Prestige* was also often defined based on Impact Factor and Metrics
16 and by the Review Process used by the journal.

17

18

1 **Table 2.** Categories identified in the definitions of Prestigious.
 2 Name, description and examples of the categories found in participants' definitions of Prestige. Numbers in square
 3 brackets represent anonymized identification of participants.

Category	Description of the category	Examples
Impact factor and metrics	<i>Established measurement for the journals by the articles that are published, and the citation generated</i>	<ul style="list-style-type: none"> ☒ Highly ranked [561] ☒ Super impact factor and circulation [761]
Value	<i>Scientific rigor and applicability of the articles published in the journal.</i>	<ul style="list-style-type: none"> • Field changing, important, correct [488] • Usually very good quality [1351]
Readership	<i>Focuses on how much the published work is read by academic and non-academic people</i>	<ul style="list-style-type: none"> • Widely read [1171] • Large reading audience [86]
Relation to associations	<i>Relation of the journal to organizations that support its operation.</i>	<ul style="list-style-type: none"> • Association sponsored [812] • Affiliated with a widely recognized organization [577]
Reputation	<i>Recognition of the journal itself, the authors of the articles published in the journal or the authors citing the journal.</i>	<ul style="list-style-type: none"> • High name recognition [581] • Held in high regard by researchers [214]
Review process	<i>Elements related to the process of reviewing the articles that are published, such as editors, feedback and rejection rate</i>	<ul style="list-style-type: none"> • Expert peer-review [388] • Hard to get accepted for publication [365]



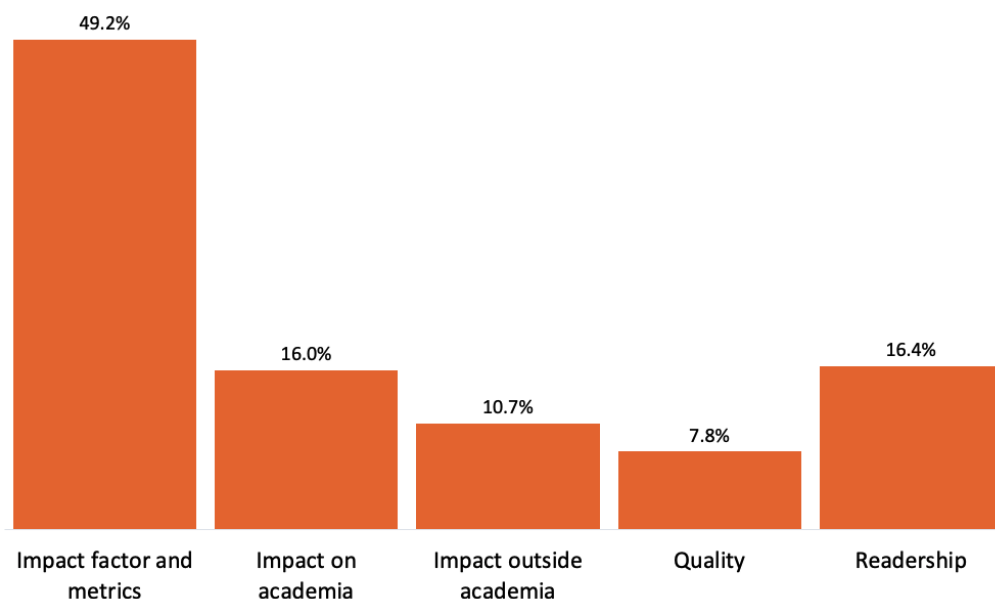
4
 5 **Figure 2.** Definitions of Prestige.
 6 Percentage of responses that contained at least one segment in each of the categories identified in participants'
 7 definitions of Prestige.

1 Finally, the analysis of the definitions provided for *high impact* resulted in 242 segments.
 2 These segments were coded under six categories: *Impact Factor and Metrics*, *Impact on*
 3 *Academia*, *Impact Outside Academia*, *Quality*, and *Readership*. Table 3 provides definitions for
 4 each of the categories, as well as some examples for each of them.

5 The result of this coding process shows that high impact is defined by the JIF or other
 6 citation metrics in almost half of all instances (Figure 3). Definitions in this category included:
 7 “some factor that assess the impact” [581], “number of citations/ papers” [478] or “Interplay
 8 between Impact Factor and number of cites per year” [500]. To a lesser extent, *High Impact* was
 9 defined by the volume of readership the research receives and by the impact that the work had
 10 on practice, public policy, or in the media.

11 **Table 3.** Categories identified in the definitions of High Impact.
 12 *Name, description and examples of the categories found in participants’ definitions of High Impact. Numbers in*
 13 *square brackets represent anonymized identification of participants.*

Category	Description of the category	Examples
<i>Impact factor and metrics</i>	Established measurement for the journals by the articles that are published, and the citation generated	<ul style="list-style-type: none"> • Interplay between Impact Factor and number of cites per year [500] • Use of impact factor to identify a journal's worthiness [862]
<i>Impact on academia</i>	Relevance and influence of the articles on future research	<ul style="list-style-type: none"> • Immediately impacting the next work to be published [583] • Influences a lot of other researchers [391]
<i>Impact outside academia</i>	Impact on practices and public policies by the articles, as well as replicability on media outlets	<ul style="list-style-type: none"> • Immediate impact on practice [1256] • Impact on policy & practice [478]
<i>Quality</i>	Scientific rigor and applicability of the articles published in the journal.	<ul style="list-style-type: none"> • Rigorous, robust, important, field changing, important, correct [488] • High quality research [732]
<i>Readership</i>	Focuses on how much the published work is read by academic and non-academic people	<ul style="list-style-type: none"> • High readership, broad readership [154] • Read widely [1171]



1

2 **Figure 3.** Definitions of High Impact.

3 *Percentage of responses that contained at least one segment in each of the categories identified in participants'*
4 *definitions of High Impact.*

5 **Differences by demographics**

6 We performed a Chi-Square test on the definitions of the three terms to understand if
7 they varied depending on the gender, age, and academic discipline of the faculty member or
8 according to the institution type to which they belong (R-type or M-type). The definitions
9 provided by surveyed faculty do not have significant variation in any of these categories, which
10 implies that the academics conceive of these terms irrespective of their gender, academic
11 discipline or type of institution in which they are employed (Table 4 and Table 5).

12

1 **Table 4.** Breakdown of definitions by demographic characteristics. *Overview of the participants'*
2 *definition of Quality, Prestige and Impact by their gender, institution type, and age.*

		Gender		Institution Type		Age				
		Male	Female	R-type	M-type	25–34	35–44	45–54	55–64	65+
N =		136	154	202	91	25	82	98	62	23
High Quality	Impact factor and metrics	14.7%	8.4%	13.4%	7.7%	8.0%	9.8%	14.3%	4.8%	30.4%
	Quality and relevance	34.6%	37.0%	35.1%	35.2%	36.0%	37.8%	32.7%	38.7%	34.8%
	Readership	2.9%	1.9%	3.0%	1.1%	4.0%	1.2%	2.0%	4.8%	0.0%
	Reputation	7.4%	9.1%	8.4%	8.8%	4.0%	7.3%	12.2%	9.7%	0.0%
	Review process	40.4%	43.5%	40.1%	47.3%	48.0%	43.9%	38.8%	41.9%	34.8%
N =		125	129	184	76	20	77	85	55	22
Prestige	Impact factor and metrics	16.8%	16.3%	19.0%	17.1%	30.0%	19.5%	18.8%	12.7%	18.2%
	Quality and relevance	13.6%	10.9%	12.0%	13.2%	10.0%	10.4%	14.1%	16.4%	4.5%
	Readership	4.8%	4.7%	4.3%	5.3%	0.0%	3.9%	7.1%	5.5%	0.0%
	Relation to associations	6.4%	2.3%	4.9%	2.6%	0.0%	1.3%	4.7%	5.5%	13.6%
	Reputation	40.0%	47.3%	46.2%	35.5%	35.0%	45.5%	36.5%	47.3%	54.5%
	Review process	18.4%	18.6%	13.6%	26.3%	25.0%	19.5%	18.8%	12.7%	9.1%
N =		119	121	165	77	18	75	74	55	22
High Impact	Impact factor and metrics	50.4%	47.1%	52.1%	42.9%	55.6%	53.3%	47.3%	43.6%	50.0%
	Impact on academia	15.1%	17.4%	15.8%	15.6%	16.7%	10.7%	17.6%	20.0%	18.2%
	Impact outside academia	10.9%	9.9%	9.1%	14.3%	5.6%	10.7%	8.1%	16.4%	9.1%
	Quality	8.4%	7.4%	9.1%	5.2%	5.6%	5.3%	6.8%	12.7%	9.1%
	Readership	15.1%	18.2%	13.9%	22.1%	16.7%	20.0%	20.3%	7.3%	13.6%

3

4

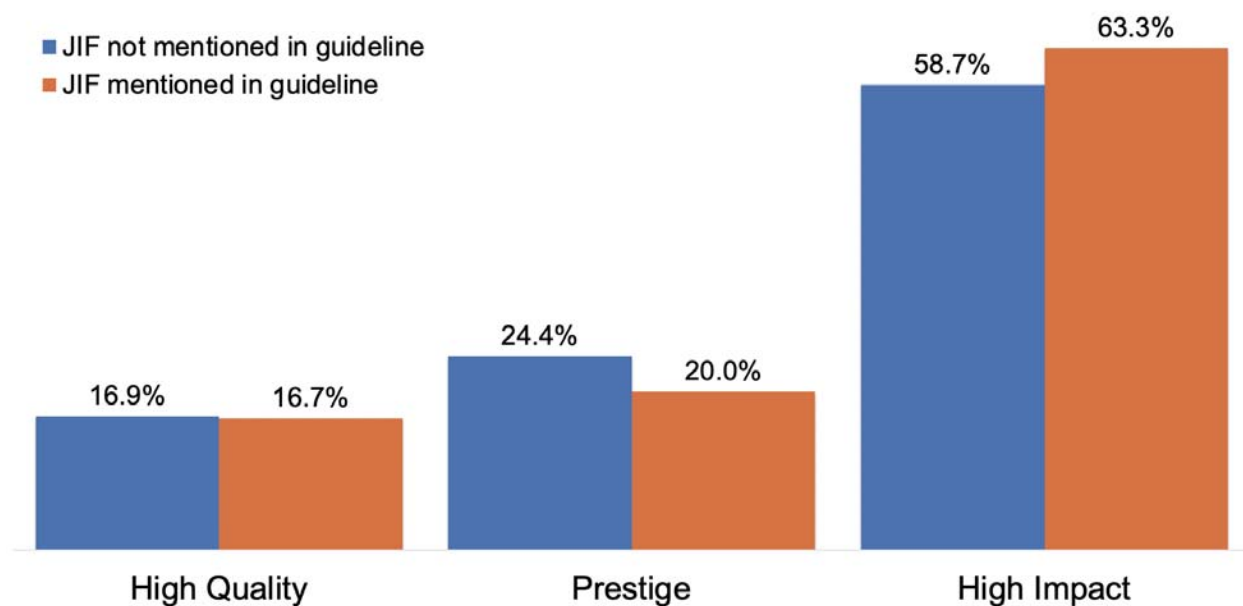
1 **Table 5.** Breakdown of definitions by demographic characteristics. *Overview of the participants'*
2 *definition of Quality, Prestige and Impact by discipline.*

		Discipline			
		Multi-disciplinary	Social Sciences and Humanities	Physical Sciences and Mathematics	Life Sciences
N =		30	142	55	63
High Quality	Impact factor and metrics	16.7%	11.3%	7.3%	14.3%
	Quality and relevance	40.0%	36.6%	32.7%	34.9%
	Readership	0.0%	2.1%	3.6%	3.2%
	Reputation	10.0%	9.2%	12.7%	3.2%
	Review process	33.3%	40.8%	43.6%	44.4%
N =		24	130	47	58
Prestige	Impact factor and metrics	16.7%	18.5%	12.8%	24.1%
	Quality and relevance	12.5%	13.1%	17.0%	6.9%
	Readership	4.2%	5.4%	4.3%	3.4%
	Relation to associations	4.2%	3.8%	6.4%	3.4%
	Reputation	45.8%	37.7%	44.7%	51.7%
	Review process	16.7%	21.5%	14.9%	10.3%
N =		22	125	45	52
High Impact	Impact factor and metrics	40.9%	51.2%	55.6%	42.3%
	Impact on academia	27.3%	12.8%	22.2%	13.5%
	Impact outside academia	18.2%	10.4%	6.7%	11.5%
	Quality	4.5%	11.2%	0.0%	7.7%
	Readership	9.1%	14.4%	15.6%	25.0%

3 Differences by RPT guidelines

4 Finally, given the importance that academics give to the JIF and other metrics to define
5 *high quality, prestige, and high impact*, we compared the responses received in the survey with
6 the RPT documents from the academic units of the respondent. In particular, we performed a
7 Chi-Square test to see if respondents used a definition coded as “Impact Factor and Metrics” for
8 each of the terms any differently if they worked at academic units that mentioned the JIF and

1 related terms in their RPT document. Figure 4 shows the prevalence of this definition among the
2 two groups of faculty (those part of academic units that mention the JIF and those who do not).
3 We do not find statistically significant differences between these groups [$\chi^2 = (5, N = 202) =$
4 $0.85, p > .05$], indicating that the mention of the JIF and related terms in RPT documents does
5 not affect how faculty define *high quality*, *high impact*, and *prestige*.



6
7 **Figure 4.** Use of "Impact Factor and Metrics" as a definition of various terms
8 *Percentage of responses that contained at least one segment in participants' definitions of High Quality, Prestige and*
9 *High Impact that relies on "Impact Factor and metrics", as a proportion of all the RPT guidelines that mentioned the*
10 *JIF and of the guidelines that did not.*

11 Discussion

12 Our analysis of how faculty define quality, impact and prestige, suggests three important
13 outcomes. First, it shows that these three terms, despite referring to three very different
14 concepts, are often defined in overlapping ways (with references to each other and to
15 themselves). Second, it shows that individual faculty members apply very different definitions to
16 each of the terms, with no single definition used by over 50% of respondents for any of the three
17 terms. Finally, the marked variance in definitions across faculty does not correspond to
18 demographic characteristics, such as the age, gender, discipline nor to characteristics of the
19 institution for which they work, including mentions of the JIF in their academic unit's RPT
20 guidelines.

1 While it is known that there is a lack of definitions for many of the terms and concepts
2 used in research assessment (Dean et al., 2016; Hatch, 2019; Moore et al., 2017; van Mil &
3 Henman, 2016), this study explores how three key terms are understood by faculty in absence
4 of these definitions. Specifically, our results indicate that the concepts of *quality*, *impact*, and
5 *prestige* are often seen as synonymous by faculty, with *high quality* being sometimes defined by
6 *reputation*, which is itself one of the most common definitions given for *prestige*; similarly, *high*
7 *impact* is at times defined by *quality*, and *high quality* is at times defined by the impact factor. In
8 fact, all three terms—*quality*, *impact*, and *prestige*—are defined through the use of the Impact
9 Factor and other citation metrics by some faculty (with the term *Impact* itself defined in this way
10 nearly half the time).

11 Given the subjective nature of these concepts and the overlapping definitions provided
12 by faculty, it is perhaps unsurprising to see that, in all three cases, some faculty resort to the
13 use of quantitative citation metrics as the basis for their definition. The rise in the use of metrics
14 in research assessment has been well documented (Aung et al., 2019; Wilsdon et al., 2015),
15 including in our own work that showed their prevalence in RPT documents and guidelines
16 (Alperin et al., 2019; McKiernan et al., 2019). In this sense, our study confirms that some faculty
17 believe that *quality*, *impact*, and *prestige* can be understood through citation metrics. However,
18 contrary to our hypothesis that faculty would be more likely to think in terms of metrics if they
19 were mentioned in the RPT guidelines for their institution, our study showed that respondents
20 were no more or less likely to use citation metrics in their definitions when this was the case. In
21 other words, despite their importance in determining individual career advancement, the
22 mention of the JIF in RPT guidelines is not correlated with how faculty define *quality*, *impact*, or
23 *prestige*.

24 This, of course, only raises further questions about how faculty arrive at their
25 understanding of *quality*, *impact*, and *prestige*. While our study does not offer direct answers to
26 these questions, it does point to the wide range of definitions that are currently used by faculty
27 when considering these career-determining aspects of research. Our study shows that in the
28 absence of common definitions, faculty are applying their own understanding of each term and
29 are doing so in ways that differ from their own colleagues. This may stem from their own
30 personal experiences with certain journals, reviewers, editors, and colleagues. For example,
31 the highest percent of respondents in our survey perceived quality as being related to the review
32 process of a journal. However, most journals do not make article reviews public, suggesting
33 that the review process of a journal is not widely known or understood by people outside their

1 own experiences. As it is widely documented that the peer review process can vary
2 significantly, this highlights how personal experiences may affect how people perceive these
3 different terms. This is precisely the situation that Hatch (2019) warns could lead to biases
4 (conscious or not) in evaluation processes and could explain in part why faculty are generally
5 unsure of what is expected of them for tenure (DeSanto & Nichols, 2017).

6 More broadly, our findings present a challenge for those seeking the most effective ways
7 to bring about research evaluation reform. Unfortunately, our initial exploration suggests that the
8 pathway for making changes to research assessment may not be as simple as clarifying how
9 definitions are presented in assessment guidelines, given that the inclusion of metrics-related
10 terms in RPT documents was not a determining factor in whether faculty used citation metrics in
11 defining *high quality*, *impact*, or *high prestige*. This is not to say that such changes would not be
12 worthwhile; guidelines, like policies and mandates, are an important way for departments and
13 institutions to signal their values, even when those guidelines are not strictly adhered to.
14 However, our research points to the need for additional cultural and environmental factors that
15 determine faculty thinking that cut across age, gender, institution types, and disciplines.

16 The reliance on metrics shown throughout this study further highlights how terms such
17 as JIF have come to be seen as a way of objectively assessing quality, impact and prestige.
18 This is problematic as it fails to account for the numerous limitations and biases that are present
19 in the creation and implementation of such metrics (Brembs et al., 2013; Haustein & Larivière,
20 2015; Sugimoto & Larivière, 2018) In this sense, it is worth recognizing how efforts towards
21 responsible metrics—advanced by organizations such as DORA, Leiden and HuMetricSS—may
22 indeed help us to move away from the most problematic uses of metrics that are laden with the
23 same challenges as ambiguous definitions.

24 Results of this study further highlight the need to more clearly implement evaluation
25 regimes that don't rely on ill-defined concepts - which we show are widely subjective and
26 variable- like those discussed here. Indeed, while we acknowledge the complexity of academic
27 careers, these findings demonstrate that when terms such as “high quality, impact and prestige”
28 are used in the context of research and journals, people perceive these differently. As such, it
29 suggests that the use of such terms may lead academics in many directions, which may result
30 in inconsistent outcomes, all of which could be judged differently by RPT committees. The
31 findings of this study invite us to reconsider how, if at all, and with what definitions, we want
32 quality, impact, and prestige to be critical components of research assessment.

References

- 1
- 2 Acker, S., & Webber, M. (2016). Discipline and publish: The tenure review process in Ontario
3 universities. In *Assembling and Governing the Higher Education Institution* (pp. 233–
4 255). Palgrave Macmillan, London. https://doi.org/10.1057/978-1-137-52261-0_13
- 5 Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, Citation Indicators, and Research
6 Quality: An Overview of Basic Concepts and Theories. *SAGE Open*, 9(1),
7 215824401982957. <https://doi.org/10.1177/2158244019829575>
- 8 Alperin, J. P., Muñoz Nieves, C., Schimanski, L. A., Fischman, G. E., Niles, M. T., & McKiernan,
9 E. C. (2019). How significant are the public dimensions of faculty work in review,
10 promotion and tenure documents? *ELife*, 8. <https://doi.org/10.7554/eLife.42254>
- 11 Alperin, J. P., Muñoz Nieves, C., Schimanski, L., McKiernan, E. C., & Niles, M. T. (2018). *Terms
12 and Concepts found in Tenure and Promotion Guidelines from the US and Canada* [Data
13 set]. <https://doi.org/10.7910/DVN/VY4TJE>
- 14 Aung, H. H., Zheng, H., Erdt, M., Aw, A. S., Sin, S. J., & Theng, Y. (2019). Investigating
15 familiarity and usage of traditional metrics and altmetrics. *Journal of the Association for
16 Information Science and Technology*, 70(8), 872–887. <https://doi.org/10.1002/asi.24162>
- 17 Belcher, B., & Palenberg, M. (2018). Outcomes and Impacts of Development Interventions:
18 Toward Conceptual Clarity. *American Journal of Evaluation*, 39(4), 478–495.
19 <https://doi.org/10.1177/1098214018765698>
- 20 Bergeron, D., Ostroff, C., Schroeder, T., & Block, C. (2014). The Dual Effects of Organizational
21 Citizenship Behavior: Relationships to Research Productivity and Career Outcomes in

- 1 Academe. *Human Performance*, 27(2), 99–128.
2 <https://doi.org/10.1080/08959285.2014.882925>
- 3 Blankstein, M., & Wolff-Eisenberg, C. (2019). *Ithaca S+R US Faculty Survey 2018* (p. 70).
- 4 Brembs, B. (2018). Prestigious Science Journals Struggle to Reach Even Average Reliability.
5 *Frontiers in Human Neuroscience*, 12. <https://doi.org/10.3389/fnhum.2018.00037>
- 6 Brembs, B., Button, K., & Munafò, M. (2013). Deep impact: Unintended consequences of journal
7 rank. *Frontiers in Human Neuroscience*, 7, 291.
8 <https://doi.org/10.3389/fnhum.2013.00291>
- 9 Cagan, R. (2013). The San Francisco Declaration on Research Assessment. *Disease Models &*
10 *Mechanisms*, 6(4), 869–870. <https://doi.org/10.1242/dmm.012955>
- 11 *Carnegie Community Engagement Classification*. (n.d.). Campus Compact. Retrieved January
12 10, 2019, from [https://compact.org/initiatives/carnegie-community-engagement-](https://compact.org/initiatives/carnegie-community-engagement-classification/)
13 [classification/](https://compact.org/initiatives/carnegie-community-engagement-classification/)
- 14 Chapman, C. A., Bicca-Marques, J. C., Calvignac-Spencer, S., Fan, P., Fashing, P. J.,
15 Gogarten, J., Guo, S., Hemingway, C. A., Leendertz, F., Li, B., Matsuda, I., Hou, R.,
16 Serio-Silva, J. C., & Chr. Stenseth, N. (2019). Games academics play and their
17 consequences: How authorship, *h*-index and journal impact factors are shaping the
18 future of academia. *Proceedings of the Royal Society B: Biological Sciences*, 286(1916),
19 20192047. <https://doi.org/10.1098/rspb.2019.2047>
- 20 Chen, C. Y. (2015). A Study Showing Research Has Been Valued over Teaching in Higher
21 Education. *Journal of the Scholarship of Teaching and Learning*, 15–32.
22 <https://doi.org/10.14434/josotl.v15i3.13319>

- 1 Dean, E., Elardo, J., Green, M., Wilson, B., & Berger, S. (2016). The importance of definitions.
2 In *Principles of Microeconomics: Scarcity and Social Provisioning*. Open Oregon.
- 3 DeSanto, D., & Nichols, A. (2017). Scholarly Metrics Baseline: A Survey of Faculty Knowledge,
4 Use, and Opinion about Scholarly Metrics. *College & Research Libraries*, 78(2), 150–
5 170. <https://doi.org/10.5860/crl.78.2.150>
- 6 DORA. (n.d.). *Good Practices – Funders – DORA*. San Francisco Declaration on Research
7 Assessment. Retrieved September 17, 2018, from [https://sfdora.org/good-](https://sfdora.org/good-practices/funders/)
8 [practices/funders/](https://sfdora.org/good-practices/funders/)
- 9 Edwards, M. A., & Roy, S. (2017). Academic Research in the 21st Century: Maintaining
10 Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition.
11 *Environmental Engineering Science*, 34(1), 51–61.
12 <https://doi.org/10.1089/ees.2016.0223>
- 13 Fox, J. W. (2012). Can blogging change how ecologists share ideas? In economics, it already
14 has. *Ideas in Ecology and Evolution*, 5(2), 74–77.
15 <https://doi.org/10.4033/iee.2012.5b.15.f>
- 16 Fraley, R. C., & Vazire, S. (2014). The N-Pact Factor: Evaluating the Quality of Empirical
17 Journals with Respect to Sample Size and Statistical Power. *PLoS ONE*, 9(10),
18 e109019. <https://doi.org/10.1371/journal.pone.0109019>
- 19 Genshaft, J., Wickert, J., Gray-Little, B., Hanson, K., Marchase, R., Schiffer, P., & Tanner, R. M.
20 (2016). Consideration of Technology Transfer in Tenure and Promotion. *Technology and*
21 *Innovation*, 197–20. <https://doi.org/10.3727/194982416X14520374943103>

- 1 Gordon, C. K. (2008). *Organizational rhetoric in the academy: Junior faculty perceptions and*
2 *roles* [University of North Texas].
3 http://digital.library.unt.edu/ark:/67531/metadc9779/m2/1/high_res_d/thesis.pdf
- 4 Gruzd, A., Staves, K., & Wilk, A. (2011). Tenure and promotion in the age of online social
5 media. *Proceedings of the American Society for Information Science and Technology*,
6 *48*(1), 1–9. <https://doi.org/10.1002/meet.2011.14504801154>
- 7 Harley, D., Acord, S. K., Earl-Novell, S., Lawrence, S., & King, C. J. (2010). Assessing the
8 future landscape of scholarly communication: An exploration of faculty values and needs
9 in seven disciplines. *Center for Studies in Higher Education*.
10 <http://escholarship.org/uc/item/15x7385g>
- 11 Hatch, A. (2019). *To fix research assessment, swap slogans for definitions*. *Nature*.
12 <https://www.nature.com/articles/d41586-019-03696-w>
- 13 Haustein, S., & Larivière, V. (2015). The use of bibliometrics for assessing research:
14 Possibilities, limitations and adverse effects. In *Incentives and Performance* (pp. 121–
15 139). Springer, Cham. https://doi.org/10.1007/978-3-319-09785-5_8
- 16 Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden
17 Manifesto for research metrics. *Nature News*, *520*(7548), 429.
18 <https://doi.org/10.1038/520429a>
- 19 Howard, J. (2013). Rise of “altmetrics” revives questions about how to measure impact of
20 research. *The Chronicle of Higher Education*, *59*(38), A6–A7.
- 21 *HuMetricsHSS Initiative*. (n.d.). <https://humetricshss.org/>

- 1 *Maclean's Rankings*. (n.d.). Retrieved September 17, 2018, from
2 <https://www.macleans.ca/education/unirankings/>
- 3 McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 276–282.
4 <https://doi.org/10.11613/BM.2012.031>
- 5 McKiernan, E. C., Schimanski, L. A., Muñoz Nieves, C., Matthias, L., Niles, M. T., & Alperin, J.
6 P. (2019). Use of the Journal Impact Factor in academic review, promotion, and tenure
7 evaluations. *ELife*, 8, e47338. <https://doi.org/10.7554/eLife.47338>
- 8 Miller, J. E., & Seldin, P. (2014). *Changing Practices in Faculty Evaluation: Can better*
9 *evaluation make a difference?* AAUP. [https://www.aaup.org/article/changing-practices-](https://www.aaup.org/article/changing-practices-faculty-evaluation)
10 [faculty-evaluation](https://www.aaup.org/article/changing-practices-faculty-evaluation)
- 11 Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., & Goodman, S. N. (2018).
12 Assessing scientists for hiring, promotion, and tenure. *PLOS Biology*, 16(3), e2004089.
13 <https://doi.org/10.1371/journal.pbio.2004089>
- 14 Moore, S., Neylon, C., Paul Eve, M., Paul O'Donnell, D., & Pattinson, D. (2017). "Excellence R
15 Us": University research and the fetishisation of excellence. *Palgrave Communications*,
16 3(1), 16105. <https://doi.org/10.1057/palcomms.2016.105>
- 17 Munafò, M. R., Stothart, G., & Flint, J. (2009). Bias in genetic association studies and impact
18 factor. *Molecular Psychiatry*, 14(2), 119–120. <https://doi.org/10.1038/mp.2008.80>
- 19 Niles, M. T., Schimanski, L. A., McKiernan, E. C., & Alperin, J. P. (2020). Why we publish where
20 we do: Faculty publishing values and their relationship to review, promotion and tenure
21 expectations. *PLOS ONE*, 15(3), e0228914.
22 <https://doi.org/10.1371/journal.pone.0228914>

- 1 Open Science Collaboration. (2015). Estimating the reproducibility of psychological science.
2 *Science*, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>
- 3 Piwowar, H. (2013). Altmetrics: Value all research products. *Nature*, 493(7431), 159–159.
- 4 Rice, D. B., Raffoul, H., Ioannidis, J. P. A., & Moher, D. (2020). Academic criteria for promotion
5 and tenure in biomedical sciences faculties: Cross sectional analysis of international
6 sample of universities. *BMJ*, m2081. <https://doi.org/10.1136/bmj.m2081>
- 7 Sanberg, P. R., Gharib, M., Harker, P. T., Kaler, E. W., Marchase, R. B., Sands, T. D., Arshadi,
8 N., & Sarkar, S. (2014). Changing the academic culture: Valuing patents and
9 commercialization toward tenure and career advancement. *Proceedings of the National*
10 *Academy of Sciences*, 111(18), 6542–6547. <https://doi.org/10.1073/pnas.1404094111>
- 11 Schimanski, L., & Alperin, J. P. (2018). The evaluation of scholarship in the academic promotion
12 and tenure process: Past, present, and future. *F1000Research*.
13 <https://doi.org/10.12688/f1000research.16493.1>
- 14 Snider, A., Hight, K., Brunson, A., Payakachat, N., & Franks, A. M. (2020). Qualitative Content
15 Analysis of Research and Scholarship Criteria within Promotion and Tenure Documents
16 of US Colleges/Schools of Pharmacy. *American Journal of Pharmaceutical Education*.
17 <https://doi.org/10.5688/ajpe7983>
- 18 Strauss, A., & Corbin, J. (1990). *Basics of Qualitative Research*. Sage publications.
- 19 Sugimoto, C. R., & Larivière, V. (2018). *Measuring research: What everyone needs to know*.
20 Oxford University Press.

- 1 Szucs, D., & Ioannidis, J. P. A. (2017). Empirical assessment of published effect sizes and
2 power in the recent cognitive neuroscience and psychology literature. *PLoS Biology*,
3 15(3), e2000797. <https://doi.org/10.1371/journal.pbio.2000797>
- 4 van Mil, J. W. F., & Henman, M. (2016). Terminology, the importance of defining. *International*
5 *Journal of Clinical Pharmacy*. <https://doi.org/10.1007/s11096-016-0294-5>
- 6 Vinyard, M., & Colvin, J. B. (2018). How research becomes impact: Librarians helping faculty
7 use scholarly metrics to select journals. *College & Undergraduate Libraries*, 25(2), 187–
8 204. <https://doi.org/10.1080/10691316.2018.1464995>
- 9 Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge,
10 S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J., & Johnson, B. (2015). *The*
11 *Metric Tide: Report of the Independent Review of the Role of Metrics in Research*
12 *Assessment and Management*. <https://doi.org/10.13140/RG.2.1.4929.1363>

13