DEFINING, MEASURING, AND REWARDING SCHOLARLY IMPACT: MIND THE LEVEL OF ANALYSIS

Ravi S. Ramani
Department of Business Administration
Earl G. Graves School of Business and Management
Morgan State University
GSBM 630
Email: ravisramaniphd@gmail.com

Herman Aguinis
Department of Management
School of Business
The George Washington University
Email: haguinis@gwu.edu

Jacqueline A-M. Coyle-Shapiro
Department of Management
London School of Economics and Political Science &
California State University San Bernardino
Email: j.a.coyle-shapiro@lse.ac.uk

DEFINING, MEASURING, AND REWARDING SCHOLARLY IMPACT:

MIND THE LEVEL OF ANALYSIS

ABSTRACT

We address the pervasive practice of using a journal's impact factor (JIF) as a proxy to assess individual researcher and article scholarly impact, and the grossly incorrect inferences this process generates. This invalid practice occurs because of confusion about the definition and measurement of impact at different levels of analysis. Specifically, JIF is a journal-level measure of impact, computed by aggregating citations of individual articles (i.e., upward effect), and is therefore not appropriate to use when measuring impact at lower levels of analysis such as that of individual researchers, or of individual articles published in a particular journal (i.e., downward effect). We illustrate the severity of the errors that occur when using JIF to evaluate individual scholarly impact, and advocate for an immediate moratorium on the exclusive use of JIF and other journal-level (i.e., higher level of analysis) measures when assessing the impact of individual researchers and individual articles (i.e., lower level of analysis). Given the importance and interest in assessing the scholarly impact of researchers and articles, we delineate measures that are level-appropriate and readily available. We discuss implications for the careers of researchers and educators and the administration and future of business schools, and conclude with actionable recommendations for internal and external stakeholders regarding the assessment of scholarly impact that can drive positive change in academic environments.

Keywords: Scholarly impact, research impact, research dissemination, research importance

摘要

我们解决了使用期刊影响因子(JIF)作为评估个体研究人员和文章的学术影响的替代品的普遍做法，以及

此过程产生的严重错误推论。这种无效做法的出现是因为在不同的分析层级上对影响的定义和衡量存在混

淆。具体来说，JIF 是一种期刊级别的影响力度量，通过汇总单个文章的引用（即向上效应）计算得出，

因此不适用于在较低的分析级别上测量影响力，例如单个研究人员或在特定期刊上发表的个别文章（即向

下效应）。我们说明了使用 JIF 评估个人学术影响时发生的错误的严重性，并主张在评估

单个研究人员和单个文章的影响时（即较低的分析层级），立即暂停仅仅使用 JIF 和其他期刊级别（即

较高的分析层级）的度量。鉴于评估研究人员和文章的学术影响的重要性和兴趣，我们描述了层级合适且

易于获得的度量。我们讨论了对研究人员和教育工作者的职业以及商学院的管理和未来的影响，并为内部

和外部的利益相关者就可以推动学术环境发生积极变化的学术影响评估提出了可行的建议。

关键词：学术影响，研究影响，研究传播，研究重要性

DEFINING, MEASURING, AND REWARDING SCHOLARLY IMPACT:
MIND THE LEVEL OF ANALYSIS

Scholarly impact is an "auditable or recordable occasion of influence arising out of research" (Haley, Page, Pitsis, Rivas, & Yu, 2017). Given the essential role of research in academia, one may assume that sophisticated indicators are used to measure its impact. However, although several measures are used in management and many other fields, the assessment of scholarly impact often relies on the Journal Impact Factor (JIF[1]; Haley, 2022; Larivière et al. 2016).

As an example of JIF's prominence, a large-scale global survey including members of the Academy of Management (AOM) found that respondents assigned more scholarly impact to publications in high- versus low-JIF outlets, and high-JIF outlet articles were more valuable for promotion & tenure (P&T; Haley et al., 2017). Moreover, 40% of respondents believed that JIF and JIF-influenced metrics, such as journal rankings, accurately captured individual scholarly impact, and an additional 32% believed it did so in particular instances (Haley et al., 2017, Table 10). Echoing these results, a survey of 129 universities in the U. S. and Canada found that 87% of them used JIF and JIF-influenced metrics to make P&T decisions (McKiernan, Schimanski, Nieves, Matthias, Niles, & Alperin, 2019). These researchers also noted that 40% of the universities explicitly mentioned the term JIF and equated it with scholarly impact. Furthermore, many business schools that we are familiar with routinely prescribe, either explicitly or implicitly, that faculty members should seek to publish in "high JIF journals" as a prerequisite

---

[1] JIF, which is produced by Clarivate's Journal Citation Report (JCR) of the Web of Science (WoS) database, is calculated as: (total number of citations from JCR year to items in "year -2" + citations from JCR year to items in "year -1") ÷ (total number of citable items in "year -2" + citable items in "year -1") (https://incites.help.clarivate.com/Content/Indicators-Handbook/ih-journal-impact-factor.htm). Citable items are "those items that comprise the figure in the denominator of the JIF calculation. These items are those identified in the Web of Science as an article, review or proceedings paper and are considered the substantive articles that contribute to the body of scholarship in a particular research field and those most likely to be cited by other articles. Other forms of journal content, such as editorial materials, letters, and meetings abstracts, are not considered as citable items." (http://help.incites.clarivate.com/incitesLiveJCR/9607-TRS/version/17).

for a positive hiring as well as promotion and tenure decision. And while P&T committees often consider several criteria, it is not uncommon, especially at the university-level where committee members are drawn from across the institution, to discuss a promotion or tenure application based primarily on the applicant's record in terms of the JIFs of the outlets in which they have published.

The evaluation of individual articles and individual researchers based on JIF is far from a mere labeling exercise. For researchers, evaluating their impact based on JIF and JIF-influenced metrics affects critical career outcomes including securing a tenure-track job, enjoying a teaching reduction to devote more time to research, obtaining additional funding (e.g., summer support, research accounts, cash bonuses), receiving a positive or negative P&T review decision, and attaining a chaired position (Abritis, McCook, & Watch, 2017; Edwards & Roy, 2017). For business schools, using JIF and JIF-influenced metrics to classify articles published by their faculty influences important outcomes such as business school rankings, fundraising, media attention, faculty recruitment efforts, and student enrollment (Aguinis, Cummings, Ramani, & Cummings, 2020; Morgeson & Nahrgang, 2008; Ryazanova, McNamara, & Aguinis, 2017).

Our article therefore is about JIF, and the *grossly incorrect inferential leap that occurs when JIF is used to assess the impact of individual articles published in that journal, and the impact of the individual researchers who have published articles in that journal*. This mistaken practice is due to confusion about the definition and measurement of impact at different levels of analysis. Specifically, JIF is a journal-level measure of impact computed by aggregating citations of individual articles (i.e., upward effect), and is therefore not appropriate for measuring impact at lower levels of analysis, such as that of individual researchers and of individual articles published in a particular journal (i.e., downward effect). Accordingly, we advocate for an

immediate moratorium on the use of JIF and other journal-level (i.e., higher level of analysis) measures to assess the impact of individual researchers and individual articles (i.e., lower level of analysis). Furthermore, we propose that this moratorium apply to other journal-level measures also incorrectly used to assess the impact of individual researchers and individual articles, such as: (a) Scimago Journal Rank (SJR; based on Scopus data, it counts citations in a given year to publications in the previous three-year publication window, weighing citations such that they are assigned a greater or lesser value based on the SJR of the journal giving the citation); (b) Source Normalized Impact per Paper (SNIP; based on Scopus data, it measures contextual citation impact by weighing citations based on the total number of citations in a subject field); (c) Article Influence Score (AIS; based on WoS data, it measures the average number of citations received by a journal's articles in the first five years after publication, and weighs citation by the quality of the journal providing the citation, normalized as a fraction of all articles in all publications); and (d) the newly released Journal Citation Indicator (JCI; based on WoS data, it is a field-normalized metric representing the average category-normalized citation impact for papers published in the prior three-year period.).

Despite using different databases, these indexes are identical in that they aggregate (i.e., average) scores from a lower (i.e., article) to a higher (i.e., journal) level of analysis. And while these practices may seem innocuous to scholars well-versed with JIF's pitfalls, treating JIF as a measure of individual impact signals that it is a prized metric to be chased, which can alter goal focus and motivation (Ordóñez, Schweitzer, Galinsky, & Bazerman, 2009) by substituting JIF-driven publications in lieu of the researcher's own interests. That is, rather than pursuing personally-meaningful work, researchers are incentivized to work on projects they believe will be attractive to high-JIF outlets. Furthermore, because the difference between tenure and losing a

tenure-track faculty position may, in some cases, be just a single high-JIF publication, using JIFs can motivate authors to engage in questionable research practices (QRPs) to meet this goal (Aguinis, Ramani, & Alabduljader, 2018).

We pause to make a few important clarifications. First, our goal is obviously not to deemphasize research. On the contrary, we want to encourage more meaningful and impactful research by helping schools develop measures tailored to their vision and mission. Second, citation-based metrics, which focus mostly on other researchers as stakeholders, capture just one dimension of research impact (Aguinis, Ramani, Alabduljader, Bailey, & Lee, 2019). Nevertheless, because of its importance and pervasive use in business schools worldwide, we restrict our examination to internal stakeholders (i.e., other researchers). Third, our article focuses on the invalid practice of using JIF to make inferences at the individual researcher and individual article level. But, despite its many shortcomings (e.g., Larivière & Sugimoto, 2019; Monastersky, 2005; Van Noorden, 2010), we recognize that using JIF may be warranted when the evaluation target is at the appropriate (i.e., journal) level. We caution, however, that even if levels of analysis are aligned, JIF should not be used in isolation. But, as we describe later, it should be used in conjunction with other indexes, additional information (e.g., measures of dispersion, graphs), and clear caveats about limitations regarding level of analysis.

The remainder of our article is organized as follows. We begin by discussing the circumstances that have contributed to the current state of affairs. Specifically, we examine how JIF originated, provide a brief review of why business schools seek to quantify scholarly impact, and discuss why JIF has become the lingua franca for scholarly impact in management and other fields. Next, we explain why using JIF (i.e., a journal-level metric) to make inferences about individual articles and researchers conflates levels of analysis, thereby leading to invalid

conclusions about scholarly impact. We use examples to illustrate these severe errors and show how reliance on JIF can over- and under-estimate the impact of individual articles and researchers. Finally, we advocate for a moratorium on the exclusive use of JIF and related journal-level measures to evaluate individual-level performance. We discuss the implications of our article for the careers of researchers and educators and the administration and future of business schools, and provide actionable recommendations for internal and external stakeholders regarding the assessment of scholarly impact that can drive positive change in academic environments.

HOW DID WE GET HERE?

The creator of JIF posited that its greatest value was "in the management of library journal collections" (Garfield, 1972: 477). That is, JIF was created to help librarians decide for which journals to purchase institutional subscriptions, and the number of issues to retain in the archives—not research evaluation. However, because it was calculated using data from the Institute for Scientific Information (ISI) (a predecessor of the current WoS), and because it provided a seemingly objective and external method to evaluate performance in an easy-to-understand format, it quickly became popular. Fast forward 50 years, and JIF is now a preeminent measure of scholarly impact within the academic community.

Importantly, although there are global variations in JIF's use and influence due to the language (i.e., English) of the works it indexes, the practice of equating individual article and researcher scholarly impact with journals' JIFs has become the norm in business schools around the world (Aguinis, Shapiro, Antonacopoulou, & Cummings, 2014; Haley, 2022; Mingers & Willmott, 2013; Ryazanova et al., 2017). Furthermore, JIF is commonly used by external agencies, either explicitly or implicitly, as a component in the evaluation of the scholarly impact

of journals and, in turn, individual scholars. For example, JIF and JIF-influenced journal rankings are utilized when making decisions about research funding in Australia, Belgium, Canada, China, Japan, Korea, and Sweden (Australian Research Council, 2021; Ebadi & Schiffauerova, 2015; Huang, Li, Zhang, & Sivertsen, 2021; Japan Society for the Promotion of Science, 2021; Jonkers & Zacharewicz, 2016; Lee, 2012; Quan, Chen, & Shu, 2017; Report, 2017). Finally, JIF-influenced journal rankings, such as those issued by the University of Texas at Dallas (UTD, 2021), the Chartered Association of Business Schools (ABS, 2021), and the Financial Times (FT-50, 2021), are used to rank business schools. While each of these practices has its flaws, our focus is specifically on the use of JIF as an indicator of the scholarly impact of articles published in a particular journal and the researchers who authored them. We turn to an analysis of these practices next.

FROM COMPARING JOURNALS TO EVALUATING INDIVIDUAL RESEARCHER AND INDIVIDUAL ARTICLE IMPACT

How did JIF "transmogrify into an evaluation of the quality of individual publications and of individual researchers" (Haley, 2022: 3)? The answer lies in the history of business schools, and confusion about the definition and measurement of impact at different levels of analysis. Business schools started as vocational and trade institutes training students to find suitable employment in industry (Bennis & O'Toole, 2005). Starting in the mid-1950s however, business schools began focusing on encouraging excellence in research as a means of gaining legitimacy within and outside the academic community (McLaren, 2019). Drawing on the scientific paradigm where research quality and productivity are critical performance and success criteria, and lacking clear-cut and objective measures to evaluate the impact of individual researchers, business schools turned to JIF (Aguinis et al., 2020; Gomez-Mejia & Balkin, 1992).

We believe the continuing ubiquity of JIF is also attributable to two additional factors: 1) need for faculty performance management systems, and 2) "basking in reflected glory" (BIRG) effect. Regarding performance management, JIF has four features that have fueled its use. First, as it is calculated by an independent organization (i.e., Clarivate's WoS), it is seemingly objective and free of personal biases. Second, because JIF includes a broad swath of mostly English-language journals, it provides a metric that can be applied to individual articles and researchers across not only management but also other business fields (e.g., entrepreneurship, organizational behavior, strategy, international business, accounting, marketing, finance, operations management). Third, using JIF clearly communicates the "rules of the game." That is, researchers know how others will judge their work, enhancing perceptions of fairness and providing well-defined goals that can motivate performance. Finally, and perhaps most importantly, using JIF to measure performance is *easy*. Evaluators can simply total up JIF of the journals a researcher has published in to gauge overall impact, and this score can be compared across researchers when making decisions about rewards such as summer support, teaching reductions, and research funding, among others (Ryazanova et al., 2017). This practice can also be extended across cohorts of researchers to evaluate candidates for P&T.

The second additional factor behind JIF's popularity lies in the "basking in reflected glory" (BIRG) effect. BIRG is the desire to associate oneself with success as a means of bolstering self-image and improving others' assessment of oneself, even if the person played little or no part in obtaining the achievement (Cialdini, Borden, Thorne, Walker, Freeman. & Sloan, 1976). That is, BIRG allows people to share in the positive evaluations and recognition of a successful venture, without personally contributing to that success. In business schools, the BIRG effect manifests in two ways. First, greater competition for students and funding means

that for many business schools, rankings are the name of the game (Aguinis et al., 2020; Morgeson & Nahrgang, 2008). An important component of these rankings is the number of articles published by individual researchers in journals listed on the UTD, FT-50, or ABS lists, each of which rely wholly or in part on JIF. Therefore, having faculty members who publish in high JIF outlets allows the business school and their deans to BIRG of these efforts. Second, in today's hypercompetitive environment where most faculty members find it difficult to publish even one article in top-JIF journals in their entire careers (Certo, Sirmon, & Brymer, 2010), publications by fellow faculty in these journals allows departmental colleagues to BIRG of the "win," leading them to implicitly accept and even encourage the use of JIF. As we discuss next however, using JIF as a measure of individual scholarly impact, whether motivated by performance management concerns or a desire to BIRG, can lead to deleterious consequences.

JIF AND INVALID INDIVIDUAL-LEVEL INFERENCES

Adopting JIF as the dominant indicator of individual scholarly impact has led business schools to fall prey to confusion about levels of analysis. Specifically, as the multilevel paradigm explains, phenomena are influenced by factors above and below the level where they reside, and failure to account for these influences leads to invalid inferences (Mathieu & Chen, 2011). Using JIF to evaluate the scholarly impact of articles and researchers is an example of this error based on a level of analysis confusion. That is, JIF is an upward effect in which citations at the lower-level (i.e., individual articles) are aggregated, and an average computed to create a metric at the higher level (i.e., journal). This calculation is correct, as it is simply the average for the higher-level construct (i.e., average journal citations; Klein & Kozlowski, 2000). The error occurs, however, when the higher-level metric of JIF is used to make inferences about the current or future scholarly impact of individual articles or researchers, a downward effect. This confusion

about levels of analysis and attendant erroneous inferences is known as the ecological fallacy (Robinson, 1950).

Confusion about levels of analysis is not restricted to JIF and individual journals, but also extends to the use of journal lists. These lists (e.g., UTD, FT-50, ABS) were developed, at least in part, by selecting outlets with high JIFs. For example, the UTD list, computed using JIFs from the late 1990s (Trieschmann, Dennis, Northcraft, & Nieme, 2000) and mostly unchanged in the last 25 years, is based on an upward effect whereby high JIF led to journal inclusion. However, level of analysis confusion means this link has become reversed, such that articles and the researchers who authored them are considered as having scholarly impact, a downward effect, based on the fact that they appear in a journal included on the list. Therefore, it is not uncommon to hear statements like "Journal X is on the FT-50 list, so researcher A's article in that journal must be impactful," or "Article B was published in Journal X, so it must be impactful."

The pernicious effects of this levels of analysis confusion are most visible in performance evaluations, especially when selecting candidates for faculty positions and evaluating researchers for P&T. While these weighty decisions are not made simplistically and exclusively based on JIF, for most business schools, research accounts for the lion's share of the evaluation (Alshare, Wenger, & Miller, 2007). Numerous empirical studies, using different types of data sources and methodological approaches, have revealed the critical role that JIF and JIF-influenced metrics play, both explicitly and implicitly, in this assessment. For example, McKiernan et al. (2019) found that 87% of the 129 universities they surveyed in the U. S. and Canada used JIF and JIF-influenced metrics to make P&T decisions, with 40% explicitly mentioning the term and equating it with scholarly impact. Similarly, Alshare et al. (2007) reported that 85% of deans of research-focused U.S. business schools accredited by the Association to Advance Collegiate

Schools of Business (AACSB) specifically looked for high-JIF publications when making P&T decisions. Also, the European Commission report on open science practices (Report, 2017) found that 68% of the 154 European universities surveyed used JIF as an indicator of scholarly impact, and Shu, Quan, Chen, Qiu, Sugimoto, and Larivière (2020) reported that universities in China commonly consider JIF when evaluating candidates for P&T. Most worryingly, Powdthavee, Riyanto, and Knetsch (2018) found that publications in lower JIF outlets are viewed so negatively—regardless of the merits of the article—that they significantly decrease the overall assessment of the researcher. Together, these data clearly document the "pervasiveness in faculty evaluations" (Haley et al., 2017: 13) of JIF and JIF-influenced metrics when evaluating individual scholarly impact.

As an illustration of the invalid inferences of using JIF to evaluate individual articles, Figure 1 presents the distribution of citations received by articles published in 2018 and 2019 and used to calculate the 2020 JIF for *Academy of Management Learning and Education* (AMLE) and seven other journals. We chose these seven additional journals because they have high JIFs, are acknowledged as influential or "A journals" on lists such as the FT-50, UTD, and ABS, and are therefore considered to have high scholarly impact.

As Figure 1 shows, citations for each journal are positively skewed (i.e., long right tail)[2]. For example, *Academy of Management Journal's* (AMJ) 2020 JIF was 10.19, but 63% of articles (i.e., 100 of 160) have been cited fewer than ten times, and just 24% (i.e., 38 of 160) accounted for at least 50% of all citations. These results are similar to those for all WoS fields combined (Larivière & Sugimoto, 2019), and show that most management articles (i.e., approximately 65%) receive far fewer citations than the JIF of the journal in which they were published, with a

---

[2] Methodological details implemented to create these graphs as well as more detailed numerical results are available from the authors upon request.

few articles (i.e., approximately 20%) accounting for at least half the journal's total citations.

To illustrate the invalid inferences of using JIF to evaluate individual scholars, we examined 81 management researchers recognized as having significant scholarly impact. These influential scholars have received the AOM's "Distinguished Award for Scholarly Contributions to Management" from 2000 to 2021 (AOM Career Award, 2021), recognition by WoS as a "Highly Cited Researcher" from 2016 to 2020 (HCR, 2021), or both.[3]

Our analysis showed that across this set of 81 influential researchers, using JIF instead of actual citations misestimates scholarly impact by an absolute value of 109%, with considerable variance in terms of over- and under-estimation of impact. These results illustrate errors that occur when, in line with the process used to calculate JIF, scholarly impact is examined using JIF over a period of just two years. However, JIF is also used to evaluate researchers over longer periods, such as, for example, when making P&T decisions. So, for each researcher, we computed the average number of citations received by their articles in three collectives frequently used to make such decisions (i.e., journals on the FT-50, UTD, and ABS lists). We also calculated the percent of each researcher's articles that received fewer citations than each researcher's own average, and the percent of articles that accounted for at least half of all of each researcher's citations. Even for the highly influential researchers in this set of, most (i.e., approximately 69%) of their articles received fewer citations than their average, with a few highly cited articles (i.e., approximately 18%) accounting for at least 50% of their total citations, with these percentages varying greatly across individual researchers.

DISCUSSION

Confusion about levels of analysis explains why using JIF to evaluate individual

---

[3] A detailed description of the procedures used to select these researchers, data collection and analysis, and complete results including tables and graphs are available from the authors upon request.

researchers and articles results in a gross over- and under-estimation of scholarly impact. This would be a problem even if the errors were small. But, as our illustrative results show, errors in inferences about individual scholarly impact are very substantial and therefore have important implications for the careers of researchers and educators and the administration and future of business schools. We describe these next.

Implications for Careers of Researchers and Educators and the Administration and Future of Business Schools

Because the dominant contemporary practice of assessing scholarly impact is to focus largely on articles in journals with higher JIFs, researchers are often faced with a "forced dichotomy" between pursuing research that counts (i.e., articles published in such journals) and research that does not (i.e., any other type of scholarly work; Aguinis et al., 2020; Harley, 2019).

In terms of articles, reliance on JIF as the arbiter of impact means that researchers are incentivized to prioritize the particular kinds of management scholarship that that may win favor at high JIF journals. There is no disputing that many articles published in journals with high JIFs are valuable and have indeed advanced knowledge in important ways. However, the excessive and myopic focus on publishing in higher JIF outlets means that most researchers will seek to author such articles, even if their personal research agenda and preferences are contrary to these conventions. We ask readers to consider: How likely is it that the typical AMLE article would find favor at *AMJ, Strategic Management Journal (SMJ), or Journal of International Business Studies (JIBS)*? Our experience suggests that it is highly unlikely. Consider, for example, the top WoS-cited articles published in AMLE in 2020: Tourish (2020), Abreu-Pederzini and Suárez-Barraza (2020), and Rousseau (2020). These articles made important contributions to management learning and education such as how we conceptualize and conduct research, the

globalization of business schools, and evidence-based management. We suggest, however, that they would not be published in an "A" publication precisely because of their content. More broadly, we believe that despite the merits of AMLE articles, high-quality research on management learning and education is unlikely to be published in the top JIF outlets.

In this environment, consider a junior researcher with an interest in management education. This researcher may want to investigate new learning theories, or pedagogical or experiential learning techniques that result in improved student outcomes. At the same time, the researcher is aware that manuscripts based on this work may not, or more accurately will not, be published in a high-JIF outlet, regardless of the study's merits. The researcher must then balance their desire for personal fulfillment and meaningfulness against the pressures of a tenure clock and the challenge of justifying their scholarly impact to a university-level P&T committee. As these committees are comprised of faculty from varied disciplines, they are likely unfamiliar with particular fields (e.g., finance faculty member evaluating someone in management), and often use JIF to gauge scholarly impact. So, while AMLE provides a high-quality outlet for management learning and education research, relying on JIF may lead committee members to conclude that it is a low-quality publication (i.e., 2021 JIF 4.37, ranked 100/381 in the Management category by Clarivate's JCR), giving the mistaken impression that such research is not impactful.

We posit that in this all-too-common scenario this researcher will choose to defer writing management education articles in favor of those that might be published in a top-JIF outlet. Furthermore, while we use management education as an illustrative example, the negative effects of using JIF to evaluate scholarly impact also extend to other domains. Stated differently, using JIF to evaluate scholarly impact results in the substitution of "publishable" articles in lieu of a

researcher's actual interests (Anderson, Elliott, & Callahan, 2021; Harley, 2019), leading to a situation where "many publications are written purely to further our careers rather than to advance knowledge" (Tourish, 2020: 99).

While the above scenario is hypothetical, evidence suggests that using JIF to evaluate scholarly impact does have real-world consequences. For example, Paulus, Rademacher, Schäfer, Müller-Pinzler, and Krach (2015: 1) used functional neuroimaging to show that the practice of using JIF to evaluate and reward scholarly impact is so ubiquitous that researchers "have incorporated the predominant reward principle of the scientific community in their reward system." In other words, JIF has literally rewired how researchers' brains respond to rewards! By prioritizing external goals and a particular kind of output (i.e., high JIF publications), these powerful signals are causing goal displacement and affecting researchers' intrinsic motivation, with potentially deleterious consequences (Chapman et al., 2019; Osterloh & Frey, 2015).

For individual researchers, the use of JIF to evaluate individual article and individual performance has three additional negative outcomes. First, using JIF severely over- or under-estimates the scholarly impact of individuals, as illustrated in our analyses. For example, for a researcher in our sample who has received both the AOM Career Award and is on the HCR list, using JIF underestimates actually accrued citations by 500%. On the other hand, for another researcher on both the AOM Career Award and HCR lists, using JIF overestimates actually accrued citations by 57%. And because JIF is based on a two-year window, using it to evaluate individual researchers can lead to widely varying results. For example, for these same two researchers, using 2019 instead of 2020 JIF leads to underestimates of 53% and 19%, respectively[4]. Therefore, using JIF to evaluate scholarly impact can make it harder for

---

[4] Results of analysis using the 2019 compared to the 2020 JIF are available from the authors upon request.

researchers to demonstrate their achievements when pursuing salient career outcomes such as job offers, positive P&T decisions, or chaired professorships.

Second, using JIF influences who becomes a management researcher and educator. Although some have voiced this concern (e.g., Burke & Rau, 2010; Harley, 2019), we posit that the negative consequences of using JIFs to assess individual scholarly impact are particularly damaging for future cohorts of business educators. Using JIF is a signal about the kind of scholarship valued by business schools, and it is possible that those who do not wish to play the "game" of chasing higher-JIF publications will opt out of pursuing doctoral degrees, or of pursuing academic careers after receiving their degrees. We believe that over time this will lead, and to some extent has already led, to reduced diversity in business schools such that practitioners wishing to return to academia will find themselves unable to compete for doctoral student positions. Even if these practitioners do return, under the current model of using JIF as a measure of scholarly impact, they are unlikely to be afforded the freedom to craft their own narrative, especially if they want to explore more "practitioner-oriented" themes, and instead be guided to pursue high JIF driven articles.

Finally, continued use of JIF as a measure of researcher scholarly impact also has negative influences for those who wish to pursue careers specifically in management education. Because this practice focuses exclusively on articles published in select academic journals, it excludes essential contributions to management education made by researchers via alternative mediums. As an illustrative example, consider the 81 highly successful and recognized individual researchers we examined. A search on Amazon.com reveals that they have authored or edited over 180 textbooks, handbooks, and edited volumes, including nine of the 29 most-commonly used textbooks in organizational behavior, human resource management, and

strategic management, as reported by Aguinis et al. (2019). The failure of a JIF-based model to recognize and reward the critical role of textbooks for the dissemination of knowledge (Cummings & Bridgman, 2016) means that researchers are less likely to devote time and other resources needed to craft these important vehicles that advance management education.

Moving Forward: Actionable Recommendations

Given the popularity of JIF, and the many errors and challenges posed by its use in evaluating individual scholarly impact, how can interested stakeholders, including university administrators, recruitment and P&T committees, funding agencies, and indeed the field as a whole, move forward? We outline three actionable recommendations.

Do not Conflate Levels of Analysis. Our first and perhaps most pressing implication is to hit "Stop" on using JIF and other journal-level of analysis metrics such as SJR, SNIP, AIS, and JCI as a primary way to assess individual article and researcher scholarly impact. We recognize that a desire for easy-to-use metrics, entrenched performance management systems, and the BIRG effect make it unlikely that JIF will be completely abandoned. Our first recommendation is therefore aimed at a stakeholder group who can provide a better-informed perspective on JIF, that is, journal editors and publishers—especially those associated with professional organizations.

Journal editors are keenly aware of the value placed on JIF. Given their interest in the metric, and easy access to journal data, we recommend that editors using JIF report, at a minimum, other descriptives and distribution properties (e.g., median, skew, kurtosis), and provide a graph of citation distributions for their journal. To facilitate interpretation, we suggest the format presented in Figure 1, which uses a common vertical scale and does not restrict or bin the number of citations (Larivière et al., 2016). Providing measures of dispersion and data

19

visualizations allows other stakeholders to see how averages, such as JIF, are not a good

representation of non-normal distributions, such as citations. For example, the graph for AMLE

presented in Figure 1 shows that one article (i.e., Kothiyal, Bell, & Clarke, 2018) accounts for

almost 7% of all citations received by the 51 articles published in AMLE in 2018–2019. Similar

trends can be seen across all the journals we examined. Because these comparisons are among

articles in the same journal, they show the scholarly impact of a focal article vis-à-vis an accurate

set of referent others. Presenting data in this manner offers a more holistic view of JIF, thereby

allowing for more accurate assessments regarding individual scholarly impact by interested

stakeholders such as business school administrators and university-level P&T committees.

We recognize that asking publishers and editors to forsake a metric that is directly

rewarding to them presents agency problems. After all, JIF was created so publishers could

signal librarians about which journals to purchase (Garfield, 1972). We reiterate that we are not

against JIF per se, but its use as a primary measure of individual-level impact. So, publishers and

editors using JIF appropriately—that is, referring to JIF as a journal-level descriptor while

providing measures of dispersion and data visualizations—is not as detrimental as applying JIF

to appraise individual researchers. Furthermore, we suggest that journals published by

professional organizations (e.g., AMLE by AOM, SMJ by Strategic Management Society, and

JIBS by Academy of International Business) can, and should, lead the way in this practice.

Use the Appropriate Level of Analysis. A second recommendation, useful for

recruitment and P&T committees, accreditation and funding agencies, and administrators, is to

use multiple measures at the appropriate level of analysis when evaluating individual scholarly

impact. That is, we do not recommend "an alternative," but "alternatives" (plural).

As a guide, Table 1 presents eight measures that can be used at the individual level of analysis. Using multiple alternatives together, as opposed to relying primarily on invalid proxies such as JIF, provides several advantages, such as an inclusive view of impact that can be tailored depending on the particular goals of the evaluation, compensating for deficiencies in any one indicator, and making it harder to "game" any one metric (Aguinis et al., 2014).

For example, consider the popular *h*-index, which is the number of publications (*h*) by a researcher, each of which has been cited at least *h* times (Hirsch, 2005). The index identifies individual influence on other researchers in terms of both number of publications and number of citations (Hirsch, 2005). However, it advantages senior faculty over those with shorter tenures, censors data (i.e., all publications receiving citations above *h* are considered similarly impactful), is context-specific (i.e., varies across fields of study), and does not account for number of years since publication (Bihari, Tripathi, & Deepak, 2021). Stakeholders interested primarily in rewarding recent performance—as opposed to older, seminal works—may therefore prioritize the contemporary *h*-index (*hc*-index), which places more value on recent publications, thereby providing data on which researchers have been most impactful in the time period under consideration (Sidiropoulos, Katsaros, & Manolopoulos, 2007).

We make two clarifications about the indicators of scholarly impact in Table 1. First, these measures do not address the "peer-review" component of P&T evaluations, which, some suggest, remains the most appropriate way to assess research and researchers. We agree that assessment of a scholar's work by qualified experts provides valuable insights and perspectives that metrics alone cannot fully capture. However, judgments of academic "worth" are inherently subjective, and because the P&T peer-review process is not double-blind, it is susceptible to unconscious biases (Cundiff, Danube, Zawadzki, & Shields, 2018; Régner, Thinus-Blanc, Netter,

Schmader, & Huguet, 2019). There is, therefore, an important role for objective measures in these evaluations. Second, the indicators in Table 1 are a guide, not an exhaustive list. Furthermore, as with most measures, the indicators themselves are not completely free of biases. For example, the *h*-index can return inconsistent results based on the evaluation period (Waltman & Van Eck, 2012). Similarly, altmetrics such as Twitter shares or Mendeley readership can be influenced by a researcher's gender, academic rank and status, and age (Sugimoto, Work, Larivière, & Haustein, 2017). Therefore, we suggest using multiple measures when evaluating scholarly impact. In addition, there are several more nuanced measures—including graphical tools using the free software *R* that require almost no coding (e.g., Haunschild, Bornmann, & Adams, 2019)—that draw on rapid advances in scientometrics and information sciences to capture different aspects of scholarly impact. Furthermore, global variations not accounted for by the content indexed by WoS (e.g., non-English language journals, coverage of specific subfields), means that these measures should also be calculated using additional databases such as Google Scholar and Scopus (Aguinis et al., 2019; Harzing & Alakangas, 2016).

We acknowledge that using multiple measures and databases imposes a cost. However, most stakeholders interested in evaluating scholarly impact already have access to the data (i.e., library subscriptions to databases). In addition, these stakeholders also likely have access to the human capital—such as graduate students or administrative assistants—who can substantially alleviate the time costs of obtaining, preparing, and even analyzing the data. Also, rapid technological advances mean that many of these steps are increasingly automated (e.g., Haunschild et al., 2019). Alternatively, researchers themselves may be tasked with providing this information. Given the many valuable outcomes associated with such evaluations, we believe that refusing to make the necessary investment is damaging for the future of the field.

Rethink Performance Management and Reward Systems. Our third suggestion is aimed

at university administrators and the need to rethink and redesign the systems used to evaluate and

reward faculty performance. This need is a direct consequence of the pervasive use of JIF and

JIF-influenced metrics when conducting these exercises. That is, because JIF and journal

classifications or rankings are inappropriate measures of individual impact, relying on them to

evaluate P&T cases, allocate rewards such as summer support or teaching load reductions, or

consider which candidate to hire for a tenure-track position are incorrect practices.

As an immediate step, we suggest that policies and procedures for performance

management systems be updated to explicitly instruct evaluators to not emphasize JIF when

judging individual researchers. While a select few have already taken this important step (e.g.,

Woolston, 2021), most universities continue to encourage the use of JIF, either explicitly (e.g.,

requiring publications in particular journals or using journal lists), or through the use of

ambiguous language (e.g., requiring publications in "high-impact" or "high-quality" outlets)

(Haley, 2022). Furthermore, and contrary to common practice in many universities, we suggest

that faculty preparing documents for yearly review or P&T evaluation be instructed to not

include data exclusively on JIF. This is important since, as our results show, a majority of a

researcher's citations and impact is derived from a minority of their articles. That is, given the

skewness of citations, relying on metrics such as JIF, which is calculated based on an assumption

of data as being normally distributed, leads to serious inferential errors about the impact of

individual researchers. Instead, we suggest that actual citation counts be provided for all

publications because citation-based metrics are based on the appropriate level of analysis (i.e.,

articles instead of journals in which they were published).

Faculty members should also be encouraged to provide data on other indicators of

impact, such as those listed in Table 1. Moreover, recent studies suggest that different altmetrics may account for global variations in how research is shared and consumed (Ortega, 2020), and even potentially as early predictors of eventual citations (Akella, Alhoori, Kondamudi, Freeman, & Zhou, 2021). Therefore, various forms of altmetrics—such as article downloads, number of reads, shares on social media (e.g., Twitter, Reddit, Facebook), and Wikipedia mentions—which accumulate much faster than citations, can be used to provide additional quantitative data for university-level P&T committees. In addition, because these altmetrics can be obtained for other articles in the same journal published in the same year (and even same issue), they provide some indication of whether a particular article is getting more traction relative to others. We caution however, that just as the number of citations does not provide information on the context or relevance of the cited article on subsequent research (e.g., cited in passing, cited as foundational argument, cited critically; Harzing, 2002), the fact that a paper is "shared" or "liked" many times does not necessarily tell us about its actual impact.

A related action that university administrators can take is to join the San Francisco Declaration on Research Assessment (DORA). DORA is an international voluntary agreement whereby signatories commit to adopting processes that provide a more accurate evaluation of scholarly impact, including abandoning the use of "journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions" (https://sfdora.org/read/). Notably, while over two thousand organizations, including journals, have already pledged to follow DORA's guidelines, only a handful of business schools and management journals have made the commitment. By joining DORA, university administrators send a clear signal about the unsuitability of JIF as a measure of individual performance.

Another step relates to the process used to make hiring decisions for recent graduates and other early-career cases. Because citations accrue slowly, the use of citation-based metrics, such as those in Table 1, may not be sufficiently discriminating when considering shorter time spans. Therefore, we suggest that evaluators develop and adopt a competency-based framework (Aguinis, 2023). As an example, consider a "high-intensity research" business school seeking to hire a tenure-track assistant professor. Besides threshold competencies (i.e., minimum qualifications) such as a completed or almost completed doctoral degree, the school could formulate differentiating competencies to distinguish between average and potentially superior performers (Aguinis, 2023). If research performance is of primary importance, these competencies might include, for example, doctoral courses taken, contribution to completed or in-progress research, expertise in particular methodologies and tools (e.g., multilevel modeling, ethnography, R), data-collection partnerships, seminars, consortia, or workshops attended, grants or external funding potential, and evaluation of the advisor's—or more broadly, dissertation committee's—research performance (Bedeian, Cavazos, Hunt, & Jauch, 2010), among others. Other competencies, perhaps less relevant for this school, such as curriculum design, teaching across different formats (e.g., in-person, online, hybrid) or audiences (e.g., undergraduate, graduate, executive), or ability to form consulting partnerships, could be weighted less. The competencies could be evaluated using absolute systems such as behavior checklists or graphic rating scales (Aguinis, 2023). Overall, the school can use customized and weighted competencies aligned with their vision and mission and reflective of the position's expectations, as opposed to relying on a coarse metric like JIF. This would help the school maintain and even enhance its distinctive research identity and reputation, while helping them avoid the trap of rewarding "A" (i.e., journal-based metrics) while hoping for "B" (i.e., individual scholar impact) (Kerr, 1975).

In addition, university administrators should fundamentally rethink career trajectories available and how performance is evaluated and rewarded within the business school. Since the rise of modern business schools about 65 years ago (McLaren, 2019), schools, the fields therein, and the depth and breadth of phenomenon examined within each field have become increasingly complex. In addition, like many knowledge-based industries, higher education is undergoing periods of rapid and sometimes disruptive change (Aguinis et al., 2014). By and large however, career trajectories within business schools have remained the same. Specifically, most business schools still primarily use three hierarchical designations to mark progress during a career (i.e., assistant, associate, and professor in the U. S., and lecturer, senior lecturer, and professor in most Commonwealth countries). Similarly, most schools still utilize the system for P&T decisions that was developed in the early to mid-1900's (AAUP, 2021), where significant career milestones are evaluated after the fifth or sixth year (i.e., for promotion from assistant to tenured associate), or between the tenth and fifteenth year (i.e., for promotion to professor). We ask readers to consider: How likely is it that career trajectories and performance management milestones used in the 1910s, and indeed even the 1970s, are valid in today's academic environment?

Our contention is that there is little similarity, be it in terms of knowledge, skills, and abilities, performance standards, teaching and service responsibilities, or work environments, between the context in which current P&T systems were developed, and the modern business school. As just one example, the time and other resources required to publish even one peer-reviewed journal article have increased dramatically over the last 20 years (Ashkanasy, 2010). Due to rising, and in many cases record-setting, number of submissions received by journals, it is not uncommon to have to wait more than two years from initial submission for an article to be accepted. Of course, this is in addition to the time required to conduct the study and develop the

manuscript in the first place. While this long timeline itself may not be damaging, it can quickly become so when paired with traditional standards which stipulate that tenure decisions must be made within five or six years of a faculty joining the school, as it may motivate researchers to engage in QRPs to meet the goal (Aguinis et al., 2018). Therefore, uncritically using these legacy systems and timelines further contributes to the performance challenges facing business schools.

Instead, we suggest that university administrators explore and adopt novel solutions that better reflect the particulars and strategic goals of their own institutions. For example, schools may amend P&T procedures by restructuring the timeline and process of evaluation. This might involve shifting from the traditional "up-or-out" process to being more flexible regarding evaluation windows, and implementing new hierarchical designations (e.g., advanced assistant, advanced associate). Such alternate career paths are particularly important in knowledge-based industries, where they can help increase employee diversity, and in turn, members of underrepresented groups in leadership positions (McGinn & Milkman, 2013). Alternatively, schools may create "expertise-based" career paths allowing faculty to progress based on excellence in a particular area, or areas, rather than overall performance. This practice is already common among knowledge-based workers in industries such as information technology, where employees can develop their careers based on technical expertise (e.g., in software development, cloud computing, or web services), instead of having to enter the managerial ranks (e.g., Brunswick, 2016; Goldsmith, 2016). In business schools, this practice might manifest as faculty having the opportunity to focus on and be rewarded for impact on one or two of the three elements (i.e., research, teaching, and service/outreach) of job performance. In fact, a few schools have already instituted some version of this program in the form of "teaching-focused faculty," whose primary focus is on student instruction.

27

For example, a scholar seeking to make their mark through management education may be more amenable to joining a school that offers an expertise-based career path that rewards the development of teaching cases, simulations, and other educational material in the same manner as journal articles. Similarly, a scholar wishing to help minority entrepreneurs start businesses may find more attractive a school that recognizes and rewards these types of outreach efforts when evaluating candidates for P&T. As another example, while many schools advertise and pay lip service to the idea of interdisciplinary work, traditional P&T policies makes such endeavors highly risky, especially for junior scholars (Else, 2017; Millar, 2013). Creating more pluralistic P&T policies to allow room for such exploration can therefore help schools recruit and employ such scholars, even when competing with more resource-rich organizations (Benson et al., 2016; Klein & Falk-Krzesinski 2017).

One area where the benefit of such innovative systems may be particularly noteworthy is in the recruitment, development, and retention of star performers. Star performers generate disproportionately greater impact over an extended period as compared to their peers, and are highly prized by organizations across industries (Groysberg, 2010). Indeed, a recently published study based on 824,924 individuals and including a broad swath of academic fields showed that adding more stars to an organization through recruitment can lead to more publications, conference presentations, grants obtained, and patents filed and received, all actions that contribute to scholarly impact (Joo, Aguinis, Lee, Kremer, & Villamor, 2021). Consider, for example, the newly-appointed dean of a business school. It is likely that among the faculty are a few stars whose impact "raises all ships" and far outpaces others, and who therefore are inordinately valuable to the school in terms of, for example, visibility, recruitment, alumni and donor outreach, or fundraising. Should the dean still build high-quality relationships with all

faculty, as suggested by traditional approaches to leadership, or focus on these impact stars? As another example, consider the relationship between two stars, or between stars and non-stars. Faculty are often expected to form broad collaborative social networks weighted towards colleagues in their primary department. However, if the goal is to drive impact, would the dean be better served by helping stars prioritize connections with other stars, regardless of expertise, such as, for example, between management and marketing, or between organizational behavior and international business? Finally, since citations and impact, even for stars, are not normally distributed, how long should the dean support these stars before expecting results? Similarly, given evidence that scientific output may be subject to diminishing returns such that, past a certain threshold, devoting more resources does not lead to greater returns (Haley, 2022), at what point should the dean stop investing in these impact stars and back others instead?

CONCLUSIONS

Defining, measuring, and rewarding scholarly impact has long been a subject of discussion within the field of management. Indeed, as Peng and Dess (2010: 282) noted, "we all write to influence and read to be influenced." A confluence of factors led business schools to adopt the JIF as a preeminent indicator of the scholarly impact and influence not just of journals (which was the original level of analysis), but of individual articles and individual researchers (which is a totally different and also unintended level of analysis). However, this journal-level metric conflates levels of analysis and is therefore not a valid indicator of the impact of individual articles and researchers. Furthermore, this metric is not valid even if we restrict our focus to researchers who have been recognized as having exceptional influence on other scholars. Accordingly, we provide three actionable recommendations that can be used to counter this challenge: 1) Do not conflate levels of analysis inherent in using JIF to judge scholarly

impact, 2) use the appropriate levels of analysis when evaluating the scholarly impact of articles and researchers, and 3) rethink performance management and reward systems. Overall, we hope that the theoretical perspectives, illustrative examples, and implications and actionable recommendations provided in our article will lead to a positive change in how we evaluate define, measure, and reward scholarly impact.

REFERENCES

AAUP. 2021. 1940 statement of principles on academic freedom and tenure. *American Association of University Professors*. Retrieved April 23, 2022 from: https://www.aaup.org/report/1940-statement-principles-academic-freedom-and-tenure

Abritis, A., McCook, A., & Watch, R. 2017. Cash bonuses for peer-reviewed papers go global. *Science*, 357, 541.

Abreu-Pederzini, G. D., & Suárez-Barraza, M. F. 2020. Just let us be: Domination, the postcolonial condition, and the global field of business schools. *Academy of Management Learning & Education*, 19: 40-58.

ABS. 2021. *Chartered Association of Business Schools Academic Journal Guide 2021*. Retrieved April 23, 2022 from: https://charteredabs.org/academic-journal-guide-2021

Aguinis, H. 2023. *Performance Management* (5th ed.). Chicago Business Press.

Aguinis, H., Cummings, C., Ramani, R. S., & Cummings, T. G. 2020. "An A is an A:" The new bottom line for valuing academic research. *Academy of Management Perspectives*, 34: 135-154.

Aguinis, H., Ramani, R. S., & Alabduljader, N. 2018. What you see is what you get? Enhancing methodological transparency in management research. *Academy of Management Annals*, 12: 83-110.

Aguinis, H., Ramani, R. S., Alabduljader, N., Bailey, J. R., & Lee, J. 2019. A pluralist conceptualization of scholarly impact in management education: Students as stakeholders. *Academy of Management Learning & Education*, 18: 11-42.

Aguinis, H., Shapiro, D. L., Antonacopoulou, E. P., & Cummings, T. G. 2014. Scholarly impact: A pluralist conceptualization. *Academy of Management Learning & Education*, 13: 623-639.

Akella, A. P., Alhoori, H., Kondamudi, P. R., Freeman, C., & Zhou, H. 2021. Early indicators of scientific impact: Predicting citations with altmetrics. *Journal of Informetrics*, 15: 101128.

Alshare, K. A., Wenger, J., & Miller, D. 2007. The role of teaching, scholarly activities, and service on tenure, promotion, and merit pay decisions: Deans' perspectives. *Academy of Educational Leadership Journal*, 11: 53-68.

Anderson, V., Elliott, C., & Callahan, J. L. 2021. Power, powerlessness, and journal ranking lists: The marginalization of fields of practice. *Academy of Management Learning & Education*, 20: 89-107.

AOM Career Award. 2021. *Career achievement awards*. Retrieved April 23, 2022 from: https://aom.org/membership/awards-and-recognition/career-achievement-awards

Ashkanasy, N. M. 2010. Publishing today is more difficult than ever. *Journal of Organizational Behavior*, 31: 1-3.

Australian Research Council. 2021. *Discovery projects*. Retrieved April 23, 2022 from: https://www.arc.gov.au/grants/grant-application/instructions-and-frequently-asked-questions

Bartneck, C., & Kokkelmans, S. 2011. Detecting *h*-index manipulation through self-citation analysis. *Scientometrics*, 87: 85-98.

Bedeian, A. G., Cavazos, D. E., Hunt, J. G., & Jauch, L. R. 2010. Doctoral degree prestige and the academic marketplace: A study of career mobility within the management discipline. *Academy of Management Learning & Education*, 9: 11-25.

Bennis, W. G., & O'Toole, J. 2005. How business schools have lost their way. *Harvard Business Review*, 83(5): 96-104.

Benson, M. H., Lippitt, C. D., Morrison, R., Cosens, B., Boll, J., Chaffin, B. C., ... & Valentin, V. 2016. Five ways to support interdisciplinary work before tenure. *Journal of Environmental Studies and Sciences*, 6: 260-267.

Bihari, A., Tripathi, S., & Deepak, A. 2021. A review on *h*-index and its alternative indices. *Journal of Information Science*. Advance online publication. https://doi.org/10.1177/01655515211014478

Brunswick. 2016. *Brunswick Corporation: Brunswick technical career path welcomes more employees*. Retrieved April 23, 2022 from: https://www.brunswick.com/news/press-releases/detail/414/brunswick-corporation-brunswick-technical-career-path

Burke, L. A., & Rau, B. 2010. The research–teaching gap in management. *Academy of Management Learning & Education*, 9: 132-143.

Certo, S. T., Sirmon, D. G., & Brymer, R. A. 2010. Competition and scholarly productivity in management: Investigating changes in scholarship from 1988 to 2008. *Academy of Management Learning & Education*, 9: 591-606.

Chapman, C. A., Bicca-Marques, J. C., Calvignac-Spencer, S., Fan, P., Fashing, P. J., Gogarten, J., ... & Chr. Stenseth, N. 2019. Games academics play and their consequences: How authorship, *h*-index and journal impact factors are shaping the future of academia. *Proceedings of the Royal Society B*, 286: 20192047.

Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. 1976. Basking in reflected glory: Three (football) field studies. *Journal of Personality & Social Psychology*, 34: 366-375.

Cummings, S., & Bridgman, T. 2016. The limits and possibilities of history: How a wider, deeper, and more engaged understanding of business history can foster innovative thinking. *Academy of Management Learning & Education*, 15: 250-267.

Cundiff, J. L., Danube, C. L., Zawadzki, M. J., & Shields, S. A. 2018. Testing an intervention for recognizing and reporting subtle gender bias in promotion and tenure decisions. *The Journal of Higher Education*, 89: 611-636.

Ebadi, A., & Schiffauerova, A. 2015. How to receive more funding for your research? Get connected to the right people!. *PloS one*, 10(7), e0133061.

Edwards, M. A., & Roy, S. 2017. Academic research in the 21st century: Maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science*, 34: 51-61.

Else, H. 2017. Junior scholars face penalties for interdisciplinary work. *Times Higher Education*. Retrieved April 23, 2022 from: https://www.timeshighereducation.com/news/junior-scholars-face-penalties-interdisciplinary-work

FT-50. 2021. *50 Journals used in FT Research Rank*. Retrieved April 23, 2022 from: https://www.ft.com/content/3405a512-5cbb-11e1-8f1f-00144feabdc0

Garfield, E. 1972. Citation analysis as a tool in journal evaluation. *Science*, 178: 471-479.

Goldsmith, K. 2016. *Building a technical career path at Spotify*. Retrieved April 23, 2022 from: https://engineering.atspotify.com/2016/02/08/technical-career-path/

Gomez-Mejia, L. R., & Balkin, D. B. 1992. Determinants of faculty pay: An agency theory perspective. *Academy of Management Journal*, 35: 921-955.

Groysberg, B. 2010. *Chasing stars: The myth of talent and the portability of performance*. Princeton, NJ: Princeton University Press.

Haley, U. C. V. 2022. *Impact and the Management Researcher*. New York, NY: Routledge.

Haley, U. C. V., Page, M. C., Pitsis, T. S., Rivas, J. Y., & Yu, K. F. 2017. *Measuring and achieving scholarly impact: A report from the Academy of Management's Practice Theme Committee*. Academy of Management, Briar Cliff, New York. Retrieved April 23, 2022 from: https://aom.org/about-aom/strategic-plan/scholarly-impact-report?terms=measuring%20and%20achieving%20scholarly%20impact

Harley, B. 2019. Confronting the crisis of confidence in management studies: Why senior scholars need to stop setting a bad example. *Academy of Management Learning & Education*, 18: 286-297.

Harzing, A. W. 2002. Are our referencing errors undermining our scholarship and credibility? The case of expatriate failure rates. *Journal of Organizational Behavior*, 23: 127-148.

Harzing, A. W., & Alakangas, S. 2016. Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics*, 106: 787-804.

Haunschild, R., Bornmann, L., & Adams, J. 2019. *R* package for producing beamplots as a preferred alternative to the *h*-index when assessing single researchers (based on downloads from Web of Science). *Scientometrics*, 120: 925-927.

HCR. 2021. *Highly Cited Researchers*. Retrieved April 23, 2022 from: https://recognition.webofscience.com/awards/highly-cited/2020/

Hirsch, J. E. 2005. An index to quantify an individual's scientific research output. *Scientometrics*, 85: 741–754.

Huang, Y., Li, R., Zhang, L., & Sivertsen, G. 2021. A comprehensive analysis of the journal evaluation system in China. *Quantitative Science Studies*, 2: 300-326.

Japan Society for the Promotion of Science. 2021. *Grants-in-aid for scientific research (KAKENHI)*. Retrieved April 23, 2022 from: https://www.jsps.go.jp/english/e-grants/index.html

Jonkers, K., & Zacharewicz, T. 2016. *Research performance based funding systems: A comparative assessment*. Luxembourg: Publications Office of the European Union. Retrieved April 23, 2022 from: https://publications.jrc.ec.europa.eu/repository/handle/JRC101043

Joo, H., Aguinis, H., Lee, J., Kremer, H., & Villamor, I. 2021. HRM's financial value from obtaining more star performers. *International Journal of Human Resource Management.* Advance online publication. https://doi.org/10.1080/09585192.2021.1948890

Kerr, S. 1975. On the folly of rewarding A, while hoping for B. *Academy of Management Journal*, 18: 769-783.

Klein, J. T., & Falk-Krzesinski, H. J. 2017. Interdisciplinary and collaborative work: Framing promotion and tenure practices and policies. *Research Policy*, 46: 1055-1061.

Klein, K. J., & Kozlowski, S. W. 2000. From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods*, 3: 211-236.

Kothiyal, N., Bell, E., & Clarke, C. 2018. Moving beyond mimicry: Developing hybrid spaces in Indian business schools. *Academy of Management Learning & Education*, 17: 137-154.

Larivière, V., Kiermer, V., MacCallum, C. J., McNutt, M., Patterson, M., Pulverer, B., ... & Curry, S. 2016. A simple proposal for the publication of journal citation distributions. *BioRxiv.* Retrieved April 23, 2022 from: https://doi.org/10.1101/062109

Larivière, V., & Sugimoto, C. R. 2019. The journal impact factor: A brief history, critique, and discussion of adverse effects. In W. Glanzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer Handbook of Science and Technology Indicators*: 3-24. Cham: Springer.

Lee, G. W. 2012. NRF in the Korean research ecosystem. ***Proceedings of the 6th ASIAHORCs Meeting.*** Beijing, China. Retrieved April 23, 2022 from: https://www.jsps.go.jp/english//asiahorcs/data/meetings/6th/korea.pdf

Mathieu, J. E., & Chen, G. 2011. The etiology of the multilevel paradigm in management research. ***Journal of Management***, 37: 610-641.

McGinn, K. L., & Milkman, K. L. 2013. Looking up and looking out: Career mobility effects of demographic similarity among professionals. ***Organization Science***, 24: 1041-1060.

McKiernan, E. C., Schimanski, L. A., Nieves, C. M., Matthias, L., Niles, M. T., & Alperin, J. P. 2019. Use of the journal impact factor in academic review promotion and tenure evaluations. ***eLife***, 2019(8): e47338.

McLaren, P. G. 2019. Stop blaming Gordon and Howell: Unpacking the complex history behind the research-based model of education. ***Academy of Management Learning & Education***, 18: 43-58.

Millar, M. M. 2013. Interdisciplinary research and the early career: The effect of interdisciplinary dissertation research on career placement and publication productivity of doctoral graduates in the sciences. ***Research Policy***, 42: 1152-1164.

Mingers, J., & Willmott, H. 2013. Taylorizing business school research: On the 'one best way' performative effects of journal ranking lists. ***Human Relations***, 66: 1051-1073.

Monastersky, R. 2005. The number that's devouring science. ***Chronicle of Higher Education***, 52(8): 14.

Morgeson, F. P., & Nahrgang, J. D. 2008. Same as it ever was: Recognizing stability in the BusinessWeek rankings. ***Academy of Management Learning & Education***, 7: 26-41.

Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. 2009. Goals gone wild: The systematic side effects of overprescribing goal setting. ***Academy of Management Perspectives***, 23: 6-16.

Ortega, J. L. 2020. Blogs and news sources coverage in altmetrics data providers: A comparative analysis by country, language, and subject. ***Scientometrics***, 122: 555-572.

Osterloh, M., & Frey, B. S. 2015. Ranking games. ***Evaluation Review***, 39: 102-129.

Paulus, F. M., Rademacher, L., Schäfer, T. A. J., Müller-Pinzler, L., & Krach, S. 2015. Journal impact factor shapes scientists' reward signal in the prospect of publication. ***PloS one***, 10(11): e0142537.

Peng, M. W., & Dess, G. G. 2010. In the spirit of scholarship. ***Academy of Management Learning & Education***, 9: 282-298.

Powdthavee, N., Riyanto, Y. E., & Knetsch, J. L. 2018. Lower-rated publications do lower academics' judgments of publication lists: Evidence from a survey experiment of economists. ***Journal of Economic Psychology***, 66: 33-44.

Quan, W., Chen, B., & Shu, F. 2017. Publish or impoverish: An investigation of the monetary reward system of science in China (1999-2016). ***Aslib Journal of Information Management,*** 69: 486–502.

Régner, I., Thinus-Blanc, C., Netter, A., Schmader, T., & Huguet, P. 2019. Committees with implicit biases promote fewer women when they do not believe gender bias exists. ***Nature Human Behavior***, 3: 1171-1179.

Report. 2017. ***Evaluation of research careers fully acknowledging open science practices***. Retrieved April 23, 2022 from: https://sfdora.org/wp-content/uploads/2020/11/os_rewards_wgreport.pdf

Robinson, W. S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review,* 15: 351–357.

Rousseau, D. M. 2020. The realist rationality of evidence-based management. *Academy of Management Learning & Education*, 19: 415-424.

Ryazanova, O., McNamara, P., & Aguinis, H. 2017. Research performance as a quality signal in international labor markets: Visibility of business schools worldwide through a global research performance system. *Journal of World Business*, 52: 831-841.

Shu, F., Quan, W., Chen, B., Qiu, J., Sugimoto, C. R., & Larivière, V. 2020. The role of Web of Science publications in China's tenure system. *Scientometrics*, 122: 1683-1695.

Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. 2007. Generalized Hirsch *h*-index for disclosing latent facts in citation networks. *Scientometrics*, 72: 253-280.

Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. 2017. Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology*, 68: 2037-2062.

Tourish, D. 2020. The triumph of nonsense in management studies. *Academy of Management Learning & Education*, 19: 99-109.

Trieschmann, J. S., Dennis, A. R., Northcraft, G. B., & Nieme Jr, A. W. 2000. Serving constituencies in business schools: MBA program versus research performance. *Academy of Management Journal*, 43: 1130-1141.

UTD. 2021. *The UTD top 100 business school research rankings*. Retrieved April 23, 2022 from: https://jindal.utdallas.edu/the-utd-top-100-business-school-research-rankings/index.php

Van Noorden, R. 2010. Metrics: A profusion of measures. *Nature News*, 465(7300): 864-866.

Waltman, L., & Van Eck, N. J. 2012. The inconsistency of the *h*-index. *Journal of the American Society for Information Science and Technology*, 63: 406-415.

Woolston, C. 2021. Impact factor abandoned by Dutch university in hiring and promotion decisions. *Nature*, 595(7867), 462-462.

Zhang, C. T. 2009. The *e*-index, complementing the *h*-index for excess citations. *PloS one*, 4(5): e5429.

TABLE 1
Indicators of Scholarly Impact at the Article and Researcher Level of Analysis

| Indicator	Brief Description	Potential Challenges in Use
Altmetrics / Plum Analytics / Mendeley / CiteULike	Tracks non-traditional sources including public policy documents, mainstream media outlets, Wikipedia, Mendeley readership, CiteUlike, course syllabi, patents, blogs, social media (e.g., Twitter, Facebook), and multimedia platforms (e.g., YouTube), among others	• Impact of older articles may not be fully captured due to data coverage issues prior to launch of website (e.g., 2012 for Altmetrics; 2011 for Plum Analytics) • Susceptible to strategic manipulation • Coverage of sources varies across different years
h-index	Represents number of publications (h) each of which have been cited at least h times	• Not normalized across fields • Does not account for: publication years of papers, number of years journal/researcher has been active, self-citations, or number of coauthors • Censors data (i.e., all publications receiving citations above h are considered similarly impactful) • Subject to strategic manipulation
hc-index	Contemporary h-index; accounts for age of articles so newer articles are more highly valued	• Not widely available • Computationally demanding[1]
aw-index	Square-root of the hc-index which allows for more recent and less-cited papers to contribute to the citation profile, even if the article does not contribute to the h-index	• Not widely available
m-index	Divides h-index by number of years since individual's first published paper to avoid disadvantaging early career researchers	• Not widely available
e-index	Represents net excess citations received by all papers outside those included in h-index calculation. Larger the e, higher the overall citation count	• Computationally demanding[2] • Not widely available
g-index	Ranks individual's article by number of citations and calculates the largest number of articles such that the top g articles receive (cumulatively) at least g^2 citations. Accounts for both number of well-cited publications and overall citation performance	• Not widely available
q-index	Indicator of strategic use of self-citations to detect possible manipulation of the h-index	• Computationally demanding[3] • Not widely available

Note: Indicators can be computed using different databases such as Web of Science, Google Scholar, and Scopus, given that they have different coverage of publications.

[1] Contemporary h-index, that is hc-index, is the number of articles in a journal or authored by a researcher such that hc of the total articles (N) have a score of $S^c(i) \geq hc$ each, and rest of articles (i.e., $N - hc$) have a score $S^c(i) \leq hc$. Formula: $S^c =(i) = \gamma * (Y(now) - Y(i) + 1)^{-\delta} * |C(i)|$, where $Y(i)$ is publication year of article i, $C(i)$ is the number of articles citing the article i, $\delta = 1$, γ = time interval chosen for measurement (Sidiropoulos, Katsaros, & Manolopoulos, 2007).

[2] Citations received by articles in a journal or authored by a researcher in excess of those included in the h-index, that is e^2, is calculated as: $e^2 = \sum_{j=1}^{h} cit_j - h^2$, where cit_j represents the citations received by the jth paper (Zhang, 2009).

[3] Strategic use of self-citations, that is q-index, is calculated as $q_{p,i} = 0$ if $i < h_p$, or $1/(i + 1 - a_{p,i} - h_p)$ if $i \geq h_p$. Further, $a_{p,i} = 0$ if $i \leq h_p$, $a_{p,i-1}$ if $i > h_p$, $c_{p,i} - c_{p,i-1} \neq 0$, and $a_{p,i-1} + 1$ if $i > h_p$, $c_{p,i} - c_{p,i-1} = 0$. $c_{p,i}$ = All author articles (p) sorted in descending order of citations, creating citation profile h_p (Bartneck & Kokkelmans, 2011).
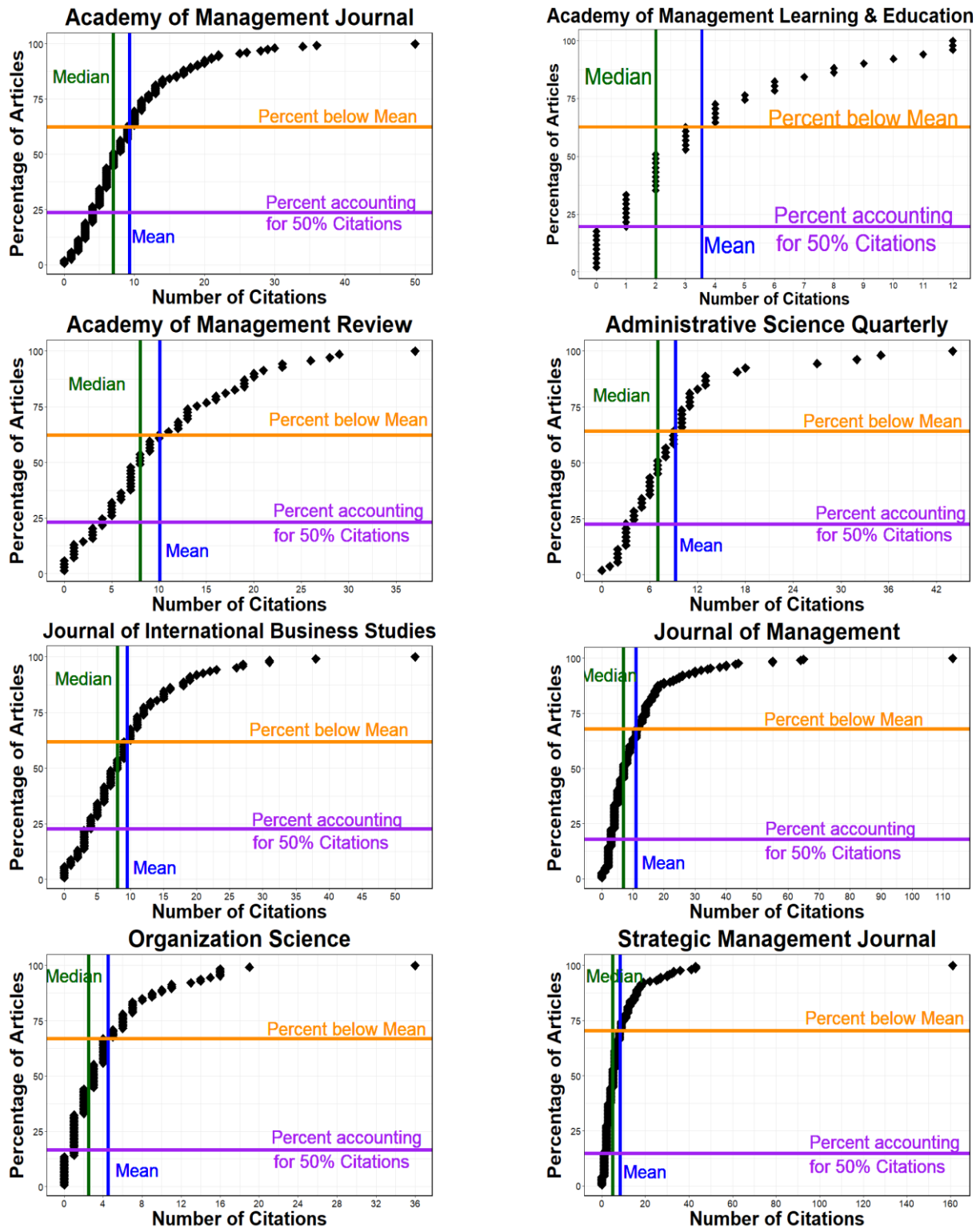
Note: Overlaid horizontal purple line shows percent of articles accounting for 50% of all citations; horizontal orange line shows percent of articles receiving fewer citations than journal's average. Overlaid vertical blue line shows mean and vertical green line shows median number of citations for articles published in each of the journals. Data as of September 19, 2021.

Author Bios

Ravi S. Ramani is an Assistant Professor of Organizational Behavior and Human Resource Management in the Earl G. Graves School of Business and Management at Morgan State University. Research interests include career development, progression, and outcomes; management education and scholarly impact; and organizational research methods. His work has received Best Paper awards from the Entrepreneurship and Management Education and Development divisions of the Academy of Management. He has published in Academy of Management Annals, Academy of Management Learning & Education, Academy of Management Perspectives, Journal of International Business Studies, Journal of Management, Organizational Research Methods, and elsewhere.

Herman Aguinis is the Avram Tucker Distinguished Scholar, Professor of Management, and Chair of the Department of Management at The George Washington University School of Business. His research addresses the acquisition and deployment of talent in organizations and organizational research methods. He published eleven books and about 200 journal articles, is a Fellow of the Academy of Management (AOM), and served as AOM President during 2021-2022. He received AOM's Research Methods Division Distinguished Career Award, AOM's Practice Theme Committee Scholar Practice Impact Award for lifetime impact on policy making and organizational practices, Society for Industrial and Organizational Psychology Scientific Contributions Award for lifetime contributions, and the Losey Award by the Society for Human Resource Management Foundation for lifetime achievement in human resource research.

Jacqueline A-M. Coyle-Shapiro is a Professor in Organizational Behavior at the London School of Economics and Political Science (LSE) and Professor in Leadership and Organizational Behavior at California State University San Bernardino. She studies psychological contracts, employment relationships and health, and social impact in poverty settings. Her work appears in journals, including Academy of Management Journal and Journal of Applied Psychology, among others. She is a Past President of the Academy of Management.