

Communications of the Association for Information Systems

Volume 45

Article 9

8-2019

Diversify Performance Metrics in Research: Thinking Within and Outside the Journal

Varun Grover

University of Arkansas, vgrover@uark.edu

Follow this and additional works at: <https://aisel.aisnet.org/cais>

Recommended Citation

Grover, V. (2019). Diversify Performance Metrics in Research: Thinking Within and Outside the Journal. *Communications of the Association for Information Systems*, 45, pp-pp. <https://doi.org/10.17705/1CAIS.04509>

This material is brought to you by the AIS Journals at AIS Electronic Library (AISeL). It has been accepted for inclusion in *Communications of the Association for Information Systems* by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.



Communications of the
Association for **I**nformation **S**ystems

Debate

DOI: 10.17705/1CAIS.04509

ISSN: 1529-3181

Diversify Performance Metrics in Research: Thinking Within and Outside the Journal

Varun Grover

Department of Information Systems

University of Arkansas

vgrover@walton.uark.edu

1 Introduction

Fitzgerald, Dennis, An, Tsutsui, and Muchala (2019) make the case that the journal impact factor (JIF), journal tiers, and, more fundamentally, journal-level metrics in general lack validity and reliability as measures of research impact. Such a claim has significant implications for disciplines such as the IS discipline in which promotion and tenure (P&T) committees routinely evaluate and promote researchers based on such metrics. For instance, a tier system of journals serves as a convenient and expeditious way to evaluate researchers' annual and tenure/promotion performance. In such a system, one can easily compare someone with four papers in an "A" journal and two in a "B" journal compared to someone with no papers in an "A" journal, five in a "B" journal, and two in a "C" journal or compared to some committee set benchmark. While such systems can be turnkey due to the need to discuss research focus, research leadership, and so on, we have reasons for committees to meet. In sum, we have got the system down—and it works to a fair degree across business disciplines at the unit and department level. Fitzgerald et al. (2019) provide evidence that any tiered journal system lacks reliability and validity when one evaluates impact with citation-based metrics due to the skewed distribution of citations that render journal rankings based on simple averages to lack congruency with journals selected through opinions. Concluding that purportedly higher-ranked journals do not have better review mechanisms for identifying impactful research has implications for our review processes as well.

Fitzgerald et al. (2019) make a compelling case and seem to recognize the limitations in their own arguments throughout the paper. However, I have two somewhat contradictory reactions to it. First, I have some concern with their approach that should at least give one pause. They provide some evidence for their case, but it is far from conclusive. For instance, citations (and citation metrics) are flawed measures. Thus, one should avoid jumping to the conclusion that we should move to paper-level assessments. Second, despite my apprehension with the evidence, I am generally sympathetic to the *conclusion* that we should explore paper-level metrics. In this rejoinder, I first discuss my concerns and then discuss points I agree with.

2 Basket Support

Researchers have conducted various studies in which they have found evidence to support the basket approach or some kind of tiered journal system. Fitzgerald et al. (2019) cite some of these papers, so I mention a few. Lowry et al. (2014) provide fairly compelling evidence that supports the tiered system. Specifically, they found that seven basket journals (from the eight in total) placed higher than all the journals they analyzed and that three in particular stood out. They used multiple metrics, including bibliographic and SNA data, to conclude the basket largely represents the best journals and that at least a two-tiered system exists. Based on a network analysis of 42 IS and non-IS journals, Jiang, Jiang, and Grover (2017) found that journal influence in the IS journal network placed *Management Information Systems Quarterly (MISQ)*, *Information Systems Research (ISR)*, and *Journal of Management Information Systems (JMIS)* as higher than the other basket journals. Similarly, Cuellar, Truex, and Takeda (2016) found an increase in citations across well-known journal sets going from lower to higher quality but argued that scholarly capital is a more complex composite of bibliographic metrics, including journal-level data. Fitzgerald et al. themselves found that the JIF metric might yield four journal groupings even though some journals may straddle across groups. Further, in most analyses, it does seem that *MISQ*, perhaps *ISR*, and, in some cases, *JMIS* represent the top journals, which concurs with the basket as a measure for quality. So, while journal-level metrics may not replicate the basket precisely, they do offer *some* value in differentiating quality.

3 Impact through Citations

Fitzgerald et al. (2019) acknowledge the many problems with citations when they refer to them as "backward-looking" measures. However, they are problematic due to how we use them. Researchers do not cite research due to its quality but rather due to how useful they find it. Highly innovative work that explores new areas will have a markedly different citation incidence than work that builds on an existing "structured" theoretical model with well-established metrics. This rationale explains why research about the technology acceptance model (TAM) receives many citations (the top two most cited papers and three from the top five most cited papers in IS focus on TAM), but do these citations reflect the model's high impact on the discipline? Research that has an existing platform (e.g., draws from existing theories), attracts much interest, is relative easy to conduct, and tends to have large citations as researchers exploit incrementalism in the red ocean (Grover & Lyytinen 2015). In contrast, more impactful, innovative research might have low

citations, (at least initially) and a more varied temporal citation pattern. On citation metrics alone, we would reward the former while shunning the latter.

Second, papers have a varied half-life. Again, Fitzgerald et al. (2019) recognize as much when discussing the two-year JIF's limitations. However, even a five-year JIF may fail to account for the average IS paper's half-life. *MISQ* lists its paper half-life as > 10 years¹, which means that only half the citations to a *MISQ* paper occur in that 10-year period. Thus, a five-year period may not reflect the paper's "true" citation impact. This point raises some doubt about the way Fitzgerald et al. (and pretty much everyone else, too) use impact factors to assess research performance.

Third, one can mold citation metrics to justify different kinds of impacts in many ways. Fitzgerald et al. (2019) use PageRank to weigh citations based on their importance to the citing paper. One could also divide citations into those that come from inside the discipline and those that come from outside it based on the argument impact outside the discipline constitutes the most important indicator of quality. Therefore, coming to the conclusion that journal-level metrics or opinion lists have no value requires more data. Perhaps the ratio of external to internal citations or some other proxy citation metric would better reproduce the basket?

4 Author Logic

Fitzgerald et al. (2019) argue that we often make the illogical error of affirming the consequent. Thus, if a high-quality paper appears in a high-quality journal, we assume that the high-quality journal has high-quality papers. Of course, that would be true if type I and type II errors did not exist in the journal review process and if we clearly assessed "quality". We could, for instance, say that, *in general*, a high-quality paper will be (or have a higher probability of being) published in a high-quality journal and, *in general*, high-quality journals have high-quality papers. Consider another logical sequence:

- Better-journals attract better papers (self-selection)
- Better journals have better editorial boards and reviewers
- Better editorial boards and reviewers conduct better review processes
- Better review processes better select papers with better theory and methods
- Better journals have papers with better theory and methods

So, if this logic is sound, Fitzgerald et al. (2019) in their analyses have found something such as: better theory and methods papers receive the same number of citations as worse theory and method papers.

If true, we need to ask why. The question raises various issues, such as whether different journals select papers that cover certain topics. It also raises issues about whether better theory and methods serves as an important consideration for having journal tiers or keeping the basket. But it does not lead one to conclude that we should abandon journal-level quality assessments.

5 Fuzzy Evidence

I am not particularly convinced by the data that Fitzgerald et al. (2019) present on journal rejection. They looked at the 50 most cited papers in *JMIS* and found that *MISQ* or *ISR* had rejected around half. Thus, these papers were "good", and *MISQ* made a type II error (rejected good papers). The fact that the most cited papers that *MISQ* rejected attracted the same number of citations that the papers it rejected attracted does not provide strong evidence for the authors' point. It shows that *MISQ* made a type II error and that *JMIS*, which researchers also consider a high-quality journal, did not. Looking at highly cited papers essentially estimates the cost of the type II error to *MISQ*. And, the *MISQ* review process likely added value (and perhaps citations) to the *JMIS* submission.

So, to reiterate my point, Fitzgerald et al. (2019) make a compelling case, but several concerns keep them from making a conclusive one.

¹ <https://misq.org/about/>

6 Personal Experience

At the 1999 ICIS in Charlotte, I participated in a panel on the impact of IS research in terms of relevance. In my presentation, I raised a hypothetical to the audience that we do away with all journals and publish our papers directly on the Internet. I argued that, for researchers, journals serve as “institutions of credibility” based on their affiliations with prestigious associations and the quality of their editorial boards. Therefore, instead of having journals as containers for papers, they can just serve as the review board and affiliation. So, submission to *MISQ* would entail obtaining a rating from *MISQ* (say three out of five stars) based on the journal’s mission and values. One could submit a paper to multiple journals and obtain ratings from each, which one could then put on the cover page to market the paper. There could even be a monetary system set up to foster reviews and a market price on paper downloads. I asked the audience: would the market price of a paper that received three stars from *MISQ* and five stars from a practitioner journal such as *Sloan Management Review* attract a higher price than one that received five stars from *ISR* and three stars from *Sloan Management Review*? Would the market price serve as a good proxy for the research’s value (impact)? Even though the Internet remained in its early stages in 1999 when I asked the question, most of the audience still felt that such a market system would eventually work. The market would set varying prices for different types of research depending on the market’s structure/size. The audience seemed to somewhat agree that such a system would take a long time to appear but that we would get there in the next 15 years. Well, today, we have open access journals, but they still do not have the traction of established journals, are often subject to high fees, and still mediated by academic publishers.

I learned from this experience that people are generally bullish about market-based systems to determine “price” and “value”, but certification or credibility mechanisms need to mediate these systems. Journals have traditions, boards, reputations, processes, mission statements—all of which add value. This value manifests when reviewing academic performance through journal tiers.

7 Putting It Together

To summarize my rejoinder to Fitzgerald et al.’s (2019) paper, I feel that, while I find the evidence that opinion lists and journal-level citation based lists do not coincide compelling, I do not find it decisive. I believe that concluding that we do not need these journal lists and should move to solely paper-level metrics (even with caveats that Fitzgerald et al. note) lacks justification at this stage given the analyses. Extending this logic, one might argue, for instance, that we do not need journals at all since they only “curate” and add no inherent value, which actually comes from authors.

I would argue that we need journals and the discipline’s collective “opinion” about journals’ quality. Opinions have more information than citations—they reflect experienced members’ general perception about a journal, which includes the quality of its backing (e.g., professional association such as the Association for Information Systems or the Institute for Operations Research and the Management Sciences), the reputational capital that it has painstakingly built and needs to defend, the excellence of its editorial board, and the quality of its review process.

Consider the following two scenarios:

- 1) Two journals—an A journal (with institutional credibility and a stellar editorial board) and a C journal (with little academic credibility and a weak editorial board)—both house paper X. Even if the paper received equivalent citations in both cases, I would reward the authors more for passing the A journal’s challenging review process. I would assume that the paper had good theory/methodology if the A journal published it but would not know whether it had good theory/methodology if the C journal published it. Accordingly, I would incorporate the journal’s reputation capital (board and institutional credibility) into the paper. If we rewarded the situations equivalently, in the limiting case researchers would likely submit their papers to the journal with the lowest standards. Rather than fighting for the top, they would fight for the bottom.
- 2) Paper X has strong theory and methodology that deals with an emergent area. Paper Y deals with an area saturated with incremental research that builds off a core platform model. Paper Y fails to get into an A journal but does get into a C journal. Paper X, however, gets accepted into the A journal. Because paper Y focuses on an area that already has a huge body of work, it receives a large number of citations in the short run (within five years). Because paper X focuses on an emergent (niche) area with novel theorizing, it does not receive many citations in the short run. If we only considered paper-level metrics, we would reward paper Y more than Paper X.

While artificial, these scenarios illustrate the core point. Fitzgerald et al. (2019) do not present convincing enough analyses to draw any strong conclusions about discarding journal-level metrics. At best, the skewed citations patterns do render averages less meaningful, but, even then, journal-level metrics do contain some valuable information. So, I would argue that we should use *journal-level metrics (including opinion lists) in conjunction with paper-level metrics*. A performance evaluation based journal lists reflects a candidate's training and ability to navigate review processes and the product's theoretical/methodological quality. It also contains some information about citations —perhaps that higher journals better reflect a longer-term impact. Paper-level metrics indicate how well the discipline has consumed the research, mostly in the short run. Collectively, both journal- and paper-level metrics can provide a complete picture of a candidate's performance.

Fitzgerald et al. (2019) make their recommendations cautiously but seem to have an inclination toward moving to paper-level metrics. Perhaps the discipline will move to such metrics, but, if so, we would somehow need to include journals' institutional credibility processes when evaluating impact. Until then, I suggest we diversify our evaluation systems to include paper-level assessments and continue to work on better journal-level metrics.

References

- Cuellar, M. J., Truex, D. P., & Takeda, H. (2016b). Can we trust journal rankings to assess article quality? In *Proceedings of the 22nd Americas Conference on Information Systems*.
- Fitzgerald, B., Dennis, A. R., An, J., Tsutsui, S., & Muchala, R. C. (2019). Information systems research: Thinking outside the basket and beyond the journal. *Communications of the Association for Information Systems*, 45, 110-133.
- Grover, V., & Lyytinen, K. (2015). New state of play in information systems research: The push to the edges. *MIS Quarterly*, 39(2), 271-296.
- Jiang, L., Jiang, D., & Grover, V. (2017). The sphere of influence of information systems journals: A longitudinal study. *Communications of the Association for Information Systems*, 41, 150-178.
- Lowry, P. B., Moody, G. D., Gaskin, J., Galletta, D., Humpherys, S., Barlow, J. B., & Wilson, D. W. (2014). Evaluating journal quality and the association for information systems Senior Scholars' journal basket via bibliometric measures: Do expert journal assessments add value? *MIS Quarterly*, 37(4), 993-1012.

About the Authors

Varun Grover is the David D. Glass Endowed Chair and Distinguished Professor of IS at the Walton School of Business, University of Arkansas. He has published extensively in the information systems field with over 400 publications, 220 of which are in major refereed journals. Over ten recent papers have ranked him among the top four researchers globally based on number of publications in the top IS journals and citation impact. He has an h-index of 86 and around 34,000 citations in Google Scholar. In 2013, Thompson Reuters recognized him as one of 100 highly cited scholars globally in all Business disciplines. He is Senior Editor for *MISQ Executive*, Editor of the *Journal of the Association for Information Systems* Section on path breaking research, and Senior Editor (Emeritus) for *MIS Quarterly*, the *Journal of the AIS*, and *Database*. He is recipient of numerous awards from USC, Clemson, AIS, Academy of Management, DSI, the OR Society, Anbar, PriceWaterhouse, among others for his research and teaching, and is a Fellow of the AIS.

Copyright © 2019 by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712 Attn: Reprints or via e-mail from publications@aisnet.org.