A Response to “Reconsidering Counting Articles in Ranked Venues (CARV) as the Appropriate Evaluation Criteria for the Advancement of Democratic Discourse in the IS Field”

Rick Anderson
University of Utah, rick.anderson@utah.edu

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

Recommended Citation

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.
A Response to “Reconsidering Counting Articles in Ranked Venues (CARV) as the Appropriate Evaluation Criteria for the Advancement of Democratic Discourse in the IS Field”

Rick Anderson
Marriott Library
University of Utah
rick.anderson@utah.edu

Abstract:

Cuellar, Truex, and Tajeda (2019) take the position that counting the number of articles published in ranked venues is an inappropriate method of evaluating the scholarly performance of faculty. They base their contention on a number of unfounded assertions and unsupported arguments, which the author details and analyzes. They propose an alternative evaluation criterion, which they call the “scholarly capital model”. In this rejoinder, I critique this model and find it wanting.

Keywords: Scientific Method, Peer Review, Journals.

This manuscript was solicited by the Department Editor for Debates, Karlheinz Kautz.
1 Introduction

In responding to Cuellar, Truex, and Takeda (2019), I take two propositions as given. A reader who does not share these two assumptions will find my arguments unconvincing at best and incoherent at worst, so it seems important to lay them out explicitly right up front.

First, although some aspects of reality are socially constructed, there also exist other aspects of reality—the characteristics and contours of which are not alterable by processes of social negotiation. Scholarship in general and science in particular substantially (though not entirely) focus on investigating and discovering the characteristics of these aspects of reality. Those who hold either that reality is a subjective construct that humans create by means of word games, or that a subjective human mind cannot reliably perceive “objective reality”, might as well stop reading now.

Second, reason constitutes a non-negotiable first principle of scholarly inquiry. Logic and reason do not only represent hegemonic tools that the powerful apply in an attempt to retain their privilege, though certainly they may be used in that way; they also constitute the only tools that we as academics and scientists can use to evaluate and make sense of the propositions about the world that we examine. To the degree that arguments depart from the rules of logic, they cease to have a meaningful bearing on scientific inquiry and on academic inquiry more broadly (though, in some humanistic fields, one can certainly create solid scholarly products using non-rational tools, such as aesthetic and creative judgment). For this reason, what we call the “scientific method” is not merely a clever game designed to preserve the privilege of those who belong to the “scientific method club”; rather, it involves applying non-negotiable and irreducible logical principles in the context of scientific inquiry. Those who hold that reason represents only one among many equally legitimate tools for making scientific sense of objective reality might as well stop reading now.

The topic at hand (i.e., applying publishing-based criteria to questions of academic evaluation) belongs to the realm of social-scientific discourse, one in which the methods of inquiry that are normally applied to the physical sciences are brought to bear on the abstract constructs of human society. In other words, in social science, we take investigation methods normally applied to things that are not socially constructed and apply them to the study of things that are, often, entirely socially constructed. This process can resemble trying to build a house using kitchen tools—or, perhaps more aptly, building a house with rhetoric. This disconnect between the things one studies and the tools used to study them is not necessarily fatal, however; although researchers can (and regularly do) misapply the scientific method to propositions for which it is not suited, they can (and regularly do) fruitfully apply the scientific method to other propositions in the social-science sphere.

When assessing a social behavior such as academic evaluation, there is not much in the way of non-negotiable physical properties that can be examined in a purely empirical, positivistic way. Nevertheless, logic and reason remain non-negotiable in this realm. False premises, or conclusions that fail to follow logically from the premises on which they depend, are no less fatal to social-scientific arguments than to arguments in any other domain, which brings us to Cuellar et al.’s (2019) rather odd article.

In their introduction, Cuellar et al. (2019) put forward a set of assumptions that they present as assertions of self-evident fact but that are neither self-evidently true nor sufficiently justified by the authors’ subsequent arguments. Each assertion raises a set of questions that we need to answer or at least clarify in order to judge their veracity. I present and discuss each of these in turn.

For an academic field to advance, scholars must be able to freely and openly exchange ideas. (p. 189)

Is this assertion actually true? If so, how “free and open” must the exchange be—and does “free and open” mean that individuals who work in the field can freely advance their findings without fear of reprisal, that the field welcomes ideas from outside its borders and ideas from inside it, or that the published expression of the ideas is freely available to all? Do we currently have a system characterized by a free and open exchange of ideas? The authors imply that we do not—does that mean that their academic field is not advancing?

Conversely, an environment that restricts and filters the flow of ideas such that researchers in the field cannot access them will restrict the field’s development. (p. 189)

But surely whether this statement is true depends entirely on what we mean by “restricts and filters”. Depending on what we mean by that phrase, surely the opposite of this statement could also be true (e.g.,
if counterfactual, unsubstantiated, or biased noise overruns the discourse in a particular field such that the signal of substantive and truthful claims is lost).

*The process of academic publishing* should be open, transparent, participative, and democratic. (p. 189)

Cuellar et al. (2019) frequently invoke the concept of democracy as if it had a self-evident meaning in the context of scholarly discourse and as if its desirability in that context were obvious and uncontested. But, in order for it to be a useful term here, we need to at least provisionally define “democracy”, and surely its definition goes a long way towards determining how desirable a “democratic discourse” might be in the realm of scholarship. In this context, does “democracy” mean that all have an equal voice regardless of their qualifications? Or does it mean that the field gives all ideas an equal hearing and evaluates them on their merits? In their conclusion, Cueller et al. finally suggest at least one characteristic of a “democratic discourse”: that it “enables scholars to fully consider all ideas” (p. 198). But this takes us back to the question of restriction and filtering: does democracy in academic publishing require that scholars “fully consider” all ideas related to the shape of the earth?

Cuellar et al. (2019) further assert that, in an environment of restriction and filtering:

*Blockages or distortions in the discourse limit scholars’ opportunity to see and/or evaluate new ideas that might help advance the field and, thus, delay scholars from recognizing their value.* (p. 189)

But, surely, it is also true that, without appropriate restriction and filtering, such delays are just as likely due to the noise factor. Just to take two examples of important filters: without some kind of filtering for coherence, basic credibility, and methodological rigor, the scientific discourse in any field runs the risk of becoming confounded by rampant nonsense. And without filtering for bias, scientific discourse runs the risk of becoming a platform for advocacy instead of scholarship.

The authors then proceed to describe “counting articles published in ranked venues” (CPVR)—or, in other words, treating the venue of publication as a proxy value for the actual merits of an article. Publishing venues are ranked, according to this construct, according to their “quality”, and quality is assumed to be achieved by means of editorial selectivity and the peer review process. The idea behind this system is that, when articles are vetted by means of peer review, and pass successfully through that process to the point of formal publication, promotion and tenure committees can then regard their publication as a reasonable indicator of scholarly quality. Peer-reviewed journals themselves can be ranked by reference to such comparative metrics as journal impact factor (JIF) and informal reputation in the field, and committees can then judge the scholarly achievements of candidates for promotion and tenure by examining the journals in which they have published. The CARV system lifts a significant burden from the committee members in that they do not need to read and carefully evaluate the qualities of every research article itself, which would be challenging for many of them anyway since not everyone on the committee will share the same niche expertise as the candidate and may not be able effectively to evaluate the pure merits of the candidate’s work.

However (the authors argue), this system is fatally flawed. They present four reasons as for why. In their view, CARV:

1. Represents at best a partial measure of scholarly output
2. Relies on an undeveloped, implicit theoretical base
3. Relies on distorted data that lead to invalid assessments, and
4. Has “deleterious performative effects on the field” (p. 193).

The fourth assertion in the above list relies on subjective judgment and interpretation, so, in the interest of space, I focus on the first three assertions, all of which are more susceptible to evaluation on the basis of objective evidence and logic.

### 1.1 CARV is a Partial Measure of Scholarly Output

Here, Cuellar et al. (2019) take exception to the general practice of judging scholarly products based largely (or entirely) on their quality rather than on such other factors as impact. They do not clearly define “impact” but point out that “the degree to which a field takes up an article’s ideas determines the article’s
impact” (p. 193). They also object that, since CARV represents a “single-measure’ approach to evaluation” (p. 194), it is only a partial measure and, thus, insufficient.

Of course, one might respond that, when it comes to quality, CARV is (though imperfect) at least a reasonably reliable indicator; one can reasonably see publication in a journal that has established a strong reputation for selecting high-quality articles by means of peer review and rigorous editorial oversight as a solid (if not perfect) indication that the author has done high-quality scholarly work. But to respond in this way would be to fall into the authors’ trap, which they spring with the next assertion.

1.2 The Concept of Quality Embedded in CARV is Undertheorized and Implicit

Cuellar et al. (2019) argue that the CARV approach “bases its deliberations on the idea that one has defined and can measure the quality of publications” (p. 194). They see this stance as “a dubious assumption at best” (p. 194) and argue instead that “researchers generally concede the fact that no theory of quality or operationalization of quality for academic literature exists” (p. 194). They present this rather remarkable assertion as something that is “generally conceded” and, to support that claim, cite three articles, none of which supports their assertion, and two of which actively contradict it. Dean et al. (2011) do not argue that no theory of quality or operationalization of quality for academic literature exists but rather that no universally accepted method exists for determining the quality of any given journal. Locke and Lowe (2002), in direct contradiction of the claim that Cuellar et al. cite them to support, actually argue that there is a pervasive view of how best to measure journal quality (one with which they disagree). As for Straub and Anderson (2010), not only do they assert that a theory and “operationization” of quality exists for academic journals, but their entire article reviews the various criteria by which researchers currently assess quality and critically evaluates the strengths and weaknesses of each one. More importantly, though, one can in fact empirically test the proposition that no theory of quality or operationalization of quality for academic literature exists, and the test is neither difficult nor expensive: it requires simply submitting an article to a peer-reviewed journal, acting as a peer reviewer, or, absent either of those opportunities, talking to a few journal editors. Anyone who has published an article or served as a peer reviewer will know that certain quality criteria are widely recognized in the academic literature. With rare exceptions, journal editors instruct reviewers to judge academic articles on such criteria as:

- Clarity and falsifiability of hypothesis
- Logical rigor of method
- Cogency of argumentation
- Parsimony of interpretation, and
- Clarity and precision of language.

Articles that claim findings not supported by evidence gathered in the course of the study, that report on the findings of a poorly-designed study, that elaborate conclusions beyond what the evidence supports, or that are too poorly written to be easily understood (or some combination of these) are generally—I might go out on a limb here and say “universally”—considered to be of low quality. High-quality studies are those that are built on falsifiable hypotheses, that control for intervening variables, that make arguments that follow logically from the evidence gathered, that refrain from making arguments that do not follow logically from that evidence, and that are clearly and cogently written. Journals gain prestige by, among other things, consistently publishing articles that adhere to these standards.

These quality criteria are not arbitrary, nor are they merely cultural; all speak directly to the fundamental logical structure of scholarly and scientific inquiry. And thus, at this point in our consideration of the present article, we must decide whether or not reason is optional. If it is—if, in other words, logic and reason are merely one construct among many equally legitimate ones available for evaluating the soundness of scholarship—then there is nothing at all wrong with suggesting that assessments of methodological and interpretive rigor do not constitute reliable determinations of quality. If, on the other hand, reason constitutes a non-negotiable first principle of scholarly inquiry, then the quality criteria I cite above represent something more than a convenient social construct: they represent irreducible elements of genuine scholarly quality.

I note, too, Cuellar et al.’s (2019) assertion that:
In perusing the literature on evaluating scholarly quality, we found that, while researchers constantly reference the construct, they use it in an implicit manner without definition. In other words, researchers assume that others know and accept what “quality” means. (p. 194)

However, this assertion is nonsense. The properties and characteristics of sound scholarship have been discussed extensively in the literature for centuries and are taught to every college undergraduate. Academia may contain some pockets of resistance to these principles, but the idea that the literature does not explicitly and widely define “quality scholarship” as scholarship that adheres to the rules of logic and reason and, therefore, to the general rules of the scientific method is ridiculous on its face.

### 1.3 CARV Relies on Systematically Distorted Data

Cuellar et al. (2019) suggest that “CARV is problematic because it relies on data that reviewers derive from a systematically distorted discourse” (p. 194). To support this argument, they adduce several points of (dubious) evidence.

First, Cuellar et al. (2019) posit that publication in highly ranked journals is not really a good indicator of quality. They support this assertion by citing several studies finding that highly ranked journals do not account for a majority of citations in their fields. Obviously, however, whether number of citations is itself a reliable indicator of quality is an interesting question, the answer to which one cannot take as self-evident. They make no effort to support this claim. They also argue that, since highly ranked venues have limited space, we cannot assume that what these journals publish represents all of the high-quality research in a field. This is surely true, but how it supports the argument that journal rankings distort quality data (given that quality data are invariably relative) is difficult to see, since it does not follow that the highest-quality research is not published in the highest-ranked journals. The authors also argue that CARV-based evidence is distorted because journals are frequently misclassified in terms of their quality (i.e., journals are regularly ranked incorrectly by influential ranking lists that purport to sort journals by quality). However, the study that Cuellar et al. (2019) cite to support this claim—Cuellar, Truex, and Takeda (2016b)—itself relies on highly problematic and insufficiently defended assumptions about what constitutes journal quality and about the proper categorization criteria. (Note, for example, that, in their attempt to “operationalize” the concept of “quality”, Cuellar et al. (2019) appeal to market criteria that have little or nothing to do with standards of scholarship.)

Second, Cuellar et al. (2019) argue that there is controversy in fields as to which journals belong in the highest ranks and, indeed, about whether or not there ought to exist lists of “endorsed” journals. The authors provide one example of such controversy, from a regional conference in their own field, but do not demonstrate that it is widespread enough to represent significant intra-field controversy—or, indeed, to undermine the validity of CARV-based measures.

Third, Cuellar et al. (2019) argue that “topical and methodological purity considerations in the review process distort the [CARV-based] data, which privileges certain types of articles” (p. 195). In particular, they assert the existence of a trend towards “strictly positivist research” in the IS field as opposed to “interpretivist research”. To support their assertion, they point to the editorial policy of a single journal—a policy that changed over 20 years ago such that the journal now regularly publishes interpretivist articles. They offer no other evidence for this supposed trend.

In short, Cuellar et al. (2019) do not support their assertion that CARV relies on systematically distorted data with either internal logic or sufficient evidence from the literature.

### 2 Concluding Remarks

Cuellar et al. (2019) suggest that the scholarly capital model (SCM) might replace CARV. The SCM:

Rather than attempt to assess quality..., proposes that one assess the concept of scholarly capital, ‘the collection of capabilities and reputational assets that a scholar brings to an organization’ (academic department, research group, or other organization based around research) (Cuellar, Takeda, Vidgen, & Truex, 2016a, p. 4). The SCM argues that scholarly capital comprises three subconstructs: ideational influence (the extent to which the field takes up a scholar’s ideas), connectedness (a scholar’s ability to draw on the scholar’s relationships with other scholars to advance the scholar’s ideas), and venue representation (the ability to publish in venues more central to the field). (p. 197, emphasis in original)
Interestingly, this model dispenses with the concept of scholarly “quality” entirely; instead, it applies a much broader and more subjective rubric of influence and networking. Instead of looking at a putatively objective criterion to judge an author’s contribution to the field (number of publications in high-reputation journals), SCM proposes to evaluate the social role that authors play in the field: what kinds of influence do their ideas have, how they collaborate with others to advance their ideas, whether they publish in centrally important venues, and so on. SCM asks not “how good is the science?” but rather “how much influence does the author have and how does the author wield it?”

This suggests a fundamental question: why is the quality of scholarly output a factor in tenure and promotion decisions at all? If our goal in assessing publication quality is to see how and to what degree the author is exerting influence in the field, then the CARV approach certainly comes up short. If, however, the CARV approach has a more narrow purpose (assessing the actual quality of an author’s research as measured by its acceptance into publishing venues with a strong reputation for selecting significant, well-designed, responsibly interpreted, carefully written articles), then SCM is clearly not a good alternative since it engages with questions of scientific quality only tangentially, if at all); skilled networkers can, after all, wield “ideational influence” on behalf of either good or bad scholarship. In other words, it would likely be much easier for a socially skilled tenure candidate to secure an influential place in the social networks of the profession without creating rigorous new research than for the same candidate to publish poor-quality science in a top-tier journal. Which does the candidate’s college or department want more—someone who is good at accumulating social capital or someone who has built a strong reputation for doing excellent research (as judged by the standard and widely accepted criteria of scholarly quality)? Of course, P&T committees will have to answer this question for themselves. I know how I (and my college) would answer it.
References


About the Authors

Rick Anderson is Associate Dean for Collections & Scholarly Communication in the J. Willard Marriott Library at the University of Utah. He has worked previously as a bibliographer for YBP, Inc., as Head Acquisitions Librarian for the University of North Carolina, Greensboro and as Director of Resource Acquisition at the University of Nevada, Reno. He serves on numerous editorial and advisory boards and is a regular contributor to the Scholarly Kitchen (the blog of the Society for Scholarly Publishing). He has served as president of the North American Serials Interest Group, and was the recipient of the HARRASSOWITZ Leadership in Library Acquisitions Award. In 2015, he was elected President of the Society for Scholarly Publishing, and he is the author of three books, including Scholarly Communication: What Everyone Needs to Know (Oxford University Press, 2018).