

Peer Review, Judgment, and Reading

KATHLEEN FITZPATRICK

When Laura Mandell and Susan Schreibman originally invited me to contribute to this cluster of essays on the evaluation of digital scholarship for tenure and promotion, I was a faculty member. I'd just undergone a successful review for promotion to full professor on the basis of an all-digital dossier, and Laura and Susan hoped that I'd write about some of the benefits that I enjoyed and challenges that I faced in the process.

In the meantime, I have moved into a new role, director of scholarly communication for the MLA, whose charge is to assist the organization and its members in thinking through the ways that new digital publishing paradigms and platforms are changing scholars' work today. I've written extensively about these changes in my forthcoming book, *Planned Obsolescence: Publishing, Technology, and the Future of the Academy*, in which I focus specifically on the degree to which scholars, reviewers, and administrators will all need to shift their perspectives and expectations as digital texts and objects become increasingly central to scholarly communication.

These transformations will not come easily. In my own promotion review, despite the fact that I was a scholar whose work took new digital modes of communication as its explicit focus, questions were raised about the relative status of online and print publications. Among these questions, not surprisingly, was the status of my digital work with respect to peer review. New York University Press had submitted my book manuscript

The author is director of scholarly communication at the MLA.

to conventional peer reviewers, but I also got their permission to post the draft online as part of an open, peer-to-peer review process. Open processes such as these are not yet the norm in literary and linguistic fields, but experiments with open review such as those conducted by *Shakespeare Quarterly* and *Postmedieval* indicate that discussion-oriented publishing platforms are giving rise to new means of assessing scholarly work.

Though such experiments have focused on the review of individual pieces of scholarship rather than on the assessment of a body of work for promotion and tenure, they are worth our consideration here, as the assessment of a body of work relies heavily on our assumptions about the review of individual pieces of that work. The knowledge that a publication contained in a junior colleague's dossier has been subjected to recognizable forms of peer review allows us to make certain baseline assumptions about its quality. But when our modes of peer review begin to change, we must consider how our assumptions about and approaches to the work contained in the dossier must change as well.

These changes highlight a tension in our promotion and tenure reviews between the need to read and evaluate material on its merits and the need for objective, independent measures of that material's value. The MLA Task Force on Evaluating Scholarship for Tenure and Promotion argued in its 2006 final report that

when new media make new forms of scholarship possible, those forms can be assessed with the same rigor used to judge scholarly quality in print media. We must have the flexibility to ensure that as new sources and instruments for knowing develop, the meaning of scholarship can expand and remain relevant to our changing times. (44)

Yet for many reviewers, digital work confounds the established markers of quality that successfully published material will present. The imprimatur of a prestigious university press, or that of a highly selective peer-reviewed journal, is absent from many, if not most, digital projects. However, there remain available to us many other means for assessing the scholarly contribution of digital work, many other forms of the independent expert assessment that we expect peer review to provide. The challenge presented to our review committees and administrations is to recognize and accept that expert assessment where it does, in fact, already exist.

We seek out objective, independent measures of scholarly contribution in tenure and promotion reviews as a means of removing personal bias from the review process. The desire for objectivity is a good thing; none of us feel that we can ever know thoroughly enough the fields in which our colleagues work to judge adequately the impact that they are having

on those fields, and so we seek outside, impartial sources of authority to assess that impact for us. This is part of the process that John Guillory has referred to as “externalization,” which attempts to move judgment out of the subjective and into the objective. The process begins by bringing in specialists in the field who are located outside the scholar’s institution, and it continues in the review’s movement from those specialists to the department, the college, and the institution as a whole. It is “a highly formalized procession from the most internal scene to the most external” (“Evaluating” 20), one that suggests an increasing degree of objectivity along the way.

And yet, as Guillory notes, the appeal to objectivity in externalized evaluation is problematic, precisely because external, objective assessment must always be “preceded by, and based on, acts of evaluation that must be internal to the discipline” (21). For objective judgment to be judgment, it must have as its basis subjective, interested, informed expert opinion. However, as the process of externalization in reviews moves forward, an increasing emphasis comes to be placed on quantifiable markers; in this fashion, “the cumulative record of publication and other numerable measures substitute for the information-rich accounts or descriptions of scholarly work by experts in the field” (20). This kind of substitution is increasingly employed even at the department level and is all but a certainty at more external locations of the tenure and promotion review process; we use key markers of authority, such as the journal or press imprimatur that is the result of the peer review process, as a shorthand of sorts, standing in for judgment. If a prestigious press decided to publish a book, the impartial authority that the press represents allows us to assume that the book must be good.

The problems with our reliance on such objective measures become clearly visible at the highest reaches of the review process, when metrics at least arguably appropriate to one discipline—journal impact factor, for instance—are applied to fields with which they do not represent a good match. But the danger of relying on objective measures in the evaluation process is greatest at more internal sites, especially in our departments: if we do not engage in informed, careful judgment of one another’s work, how can we expect those more external to our disciplines to do so? That informed, careful judgment must of necessity include direct engagement with the work itself, leading to a great deal of inconvenient and sometimes difficult reading. Guillory notes that “[a]t the most external site of evaluation, no reading of a candidate’s work need be done at all and would even in some ways be undesirable” (20). The misunderstandings that might result from a natural scientist’s reading of literary scholarship, for instance, should perhaps be avoided. But at more internal sites of evaluation, we must reconsider our deferrals of judgment, our desire for external

markers of authority, and accept our responsibility to read the work of our colleagues.

This responsibility surfaces in no small part because there are new forms of scholarship and new modes of communication for which the traditional markers of expert assessment—press or journal imprimatur—are not so readily visible or easily grasped. Digitally published projects are often advised, reviewed, and commented on by experts in the field, and yet that oversight is often conducted in the open, frequently after the material has been released, and usually it does not result in “publication” by a neutral third-party organization. Instead, such projects, even where published under the auspices of respected university or disciplinary organizations, are tainted by association with the self-publishing that ostensibly pervades the Internet, where anyone can post anything; open, postpublication review too often casts doubt on a project’s quality. But open review produces a wide range of comprehensible, if not traditional, markers of authority, which we must learn to read alongside digitally published texts. Such markers include comments and inbound links from other Web sites, which often reveal meaningful peer engagement with digital scholarship and indicate authors’ ongoing participation in and impact on their scholarly community.

The number of such comments and links might say something about the degree of authors’ participation and impact, but there are of course good reasons for us to resist the simplistic use of such metrics to establish the authority of work published online: our colleagues in the natural sciences, not to mention those in all fields in the United Kingdom, could tell us at length about the risk of such quantified assessment formats. It’s nonetheless worth comparing these markers with the imprimatur conveyed by traditional publication. However reductive such numerical measures of a digital text’s reception may appear, the measure we currently rely on in print—the presence of a press logo—is even more reductive. Typically, when a reputable scholarly press publishes a book, we can assume that the editor and two to three independent external reviewers have read and commented on the text and that their responses have helped the text arrive at its present state. For texts published in an active digital network, however, we have the opportunity to see how many more readers respond to the author’s work, collectively producing more assessment, not less. Even more important, who those readers are and what they have to say about the text become available for assessment as well.

In other words, digital publication can provide a superabundance of critical response rather than an absence of assessment. A review committee’s objection that it doesn’t know how to determine the authority of a

project published digitally—particularly in a case in which links and comments have been gathered and presented to the evaluators—misses the mark. The how is clear: read the project, read the comments, read the articles and posts that link to it.

Of course, the suggestion to just read it presents two immediate challenges. The first is that as it is too few of us read our colleagues' work during the review process; given the pressures on our time and the feeling that we might not be qualified to judge the work itself, we prefer to rely on imprimatur and external evaluations. Guillory has pointed out in "How Scholars Read" that the vast proliferation of scholarly publication since the Cold War has made it all but impossible to keep up with reading the scholarship in our own fields today, with the result that most such work is skimmed or "read from the peripheral matter inward, from the table of contents, the index, the notes, the introduction and conclusion, then to the chapters themselves, some of which might be read closely, others scanned, others skipped altogether" (14). The difficulties we already face in keeping up with the reading we need to do for our own classes and projects are multiplied in our engagements with the work of colleagues outside our areas of specialization; it's little wonder that we often rely on others' opinions of that work.

The second challenge is that a digitally published work can result in even more material to be read, as what needs to be assessed is not just the project but also the community's engagement with it. In evaluating work on which an author's scholarly community has commented, it's necessary, to get a sense of how that community has responded, to read not just the work but the comments as well. We must not lose sight of what we're evaluating in these reviews: impact. At the moment of the tenure review, we are meant to assess not whether the candidate has reached some minimum threshold of productivity but whether the work that the candidate has done holds the promise of more such future work; we are meant to assess the contribution that this work seems poised to make to the field. In the review for promotion to full professor, we expect to see this promise borne out, resulting in a senior scholar whose work is having a clear influence. Such influence cannot be quantified, and it certainly cannot be reduced to the imprimatur of the university press with which the scholar has published. Influence can only be read in the interactions of scholars with the work.

Not all interactions are adulatory; nor should they be. In reading the evidence of how scholars respond to digital projects, whether through the comments they leave on those projects or through the online writing linked to them, we must expect to find evidence of real intellectual

engagement—engagement that sometimes takes the form of critique or disagreement. As Guillory has pointed out, the

paradox of institutional objectivity . . . is something we would probably all concede, because we know that when we look into the black box, when we look on the immanent scene of evaluation (not just in the context of tenure but in any context of evaluation), we discover there considerably more disagreement among experts within disciplines and fields than is recognized from the external perspective, from the perspective of summing opinions. (‘‘Evaluating’’ 21)

If anything, disagreement is heightened in online discourse, because agreement—without critique, without addition—is likely to go unremarked, except perhaps as a recommended link on *Twitter* or another online forum. Those evaluating online work as part of tenure and promotion cases thus must set aside the expectation of unadulterated, hyperbolic praise and instead train themselves to interpret and understand the ‘‘immanent scene of evaluation’’ that these online exchanges represent, in which critique and disagreement often indicate a level of deep interest in and involvement with the work.

In the end, I am suggesting we commit ourselves not just to read the digital work for which our colleagues are being evaluated but also to engage with the plentiful expert assessment—much of which takes the form of online discussion and peer-to-peer review—that accompanies these texts. We may not each possess sufficient knowledge to perform an adequate evaluation of our colleagues’ projects without the assistance of experts in the field—but that expert opinion, in digital publishing, is readily available, if only we’re willing to read it. We must be willing to engage in the act of judgment ourselves. Reading and interpreting are, after all, what we do; they are the very basis of scholarship in the humanities. Externalizing our judgment by deferring our authority to others and appealing to objective measures of value, in the long run, can only devalue all our work.

WORKS CITED

- Fitzpatrick, Kathleen. *Planned Obsolescence: Publishing, Technology, and the Future of the Academy*. New York: New York UP, 2011. Print.
- Guillory, John. ‘‘Evaluating Scholarship in the Humanities: Principles and Procedures.’’ *ADE Bulletin* 137 (2005): 18–33. Print.
- . ‘‘How Scholars Read.’’ *ADE Bulletin* 146 (2008): 8–17. Print.
- Report of the MLA Task Force on Evaluating Scholarship for Tenure and Promotion*. *Modern Language Association*. MLA, 2006. Web. 12 Sept. 2011.