

It is a policy change most useful to the proper evaluation of candidates for promotion and tenure. Equally as important, however, is the example we set to create processes to encourage the deliberate, detailed, fair, and wise evaluation of our disciplinary colleagues for younger faculty to emulate. ■

**RESPONSE TO SPOTLIGHT ON PROMOTION LETTERS:
FIRST, TRUST OUR OWN JUDGEMENT**

Robert C. Lieberman, *Johns Hopkins University*

DOI: 10.1017/S1049096518002184

Kurt Weyland raises an important set of issues in his provocative and well-reasoned article. Granting (or denying) tenure is one of the most important decisions that colleges and universities make. Tenure entails both a major commitment of resources over a long time and, perhaps more important, a commitment to individual teachers and scholars on the expectation that they will continue, over the remainder of their careers, to be productive and innovative scholars, effective teachers, and committed institutional citizens. In short, it is a high-stakes bet, and we are right to be concerned, along with Weyland, about whether the procedures we follow at our own institutions and across the profession consistently provide information that we need to make reasoned critical judgments. Before directly addressing Weyland's proposal for improving what he suggests is a broken process, it is worth pausing to set the challenges of the tenure system in a broader institutional context.

What should the standard be for granting tenure? In my view, the standard is simple to articulate: successful candidates for tenure should be, above all, emerging leaders in their field of scholarship.

Naturally, an institution wants to reduce risk and uncertainty in making a tenure commitment. Thus, evaluations must be not only fair but also rigorous. Tenure review remains a bedrock process of faculty governance and it relies on a familiar system of peer review. However, it is framed by a strong institutional interest in avoiding Type I errors (false positives). It is better, from a university's point of view, to deny tenure to those who go on to be "stars" in the field than to give lifetime contracts to those who turn out to be less-productive colleagues in the mature phase of their careers. (Note: I write from the perspective of highly competitive research universities; however, I think my general points apply with some adaptation to other types of institutions.)

What should the standard be for granting tenure? In my view, the standard is simple to articulate: successful candidates for tenure should be, above all, emerging leaders in their field of scholarship. However, this straightforward standard proves to be fiendishly difficult to implement. The standards of accomplishment for intellectual leadership tend to be difficult to articulate, especially in a heterogeneous field such as political science. Should we place more weight on books or articles? How do we evaluate an individual's contribution to team projects, especially as coauthorship becomes a more widely practiced norm in parts of the field? How much weight should we put on quantity of scholarly output as

opposed to assessment of quality? How do we measure intellectual influence and impact? There are no "cookie-cutter" answers to these questions that easily separate strong from weak cases; for this reason, I think it is not generally wise for institutions to write into policy precise quantitative standards for tenure. Finally, it is difficult to dispassionately evaluate colleagues who, in many cases, have become friends; they are our office neighbors, lunch partners, workout buddies, and fellow preschool parents. Those human relationships are difficult to set aside in the interests of cold professional judgments.

These considerations suggest several reasons why, as Weyland rightly notes, external-review letters play such an important role in tenure evaluations. Of course, the foundation of any tenure case must be the department's careful assessment of the candidate's record: scholarship (both quality and impact) as well as teaching and service. However, along with the department's own evaluation, external-review letters have several important roles. First, as the standard I previously articulated suggests, tenure is as much an external as an internal process. Recognition by colleagues in the profession as an emerging important voice in a set of important scholarly debates is an essential ingredient of a strong tenure case, and this is a view of the case that external-review letters uniquely provide. They are without question the best way to assess the impact and influence (or lack thereof) of individual candidates' work and their prominence in the scholarly landscape. (Quantitative measures of influence, such as citation counts and h-indices, are useful but limited indicators; they are no substitute for the careful and nuanced evaluation of expert members of the discipline.) Moreover, external-review letters can serve

as a check on the human tendency of departmental colleagues to be partial toward those we know well. As long as we exclude interested referees (i.e., those with personal or professional stakes in the outcome of the case, such as research collaborators and former teachers), we should be able to rely on external evaluations to eliminate familiarity bias from departmental decision making.

These observations about the place of external-review letters in a well-functioning tenure-review process bring us to Weyland's central claim about the problem with current external-evaluation practices: most letters appear to be positive and do not generally seem to offer a truly candid or critical assessment of a candidate's role. He is correct about this. It is rare to see an explicitly negative letter in a tenure file; most letters come in shades of positive. (I can report, at least anecdotally, from my experience as a university administrator that this phenomenon is widespread across disciplines; political science is not distinctive in this regard.) However, we should not be so quick to infer from this pattern of positivity that letters do not carry useful information about candidates. When read carefully, they reveal a great deal.

There is no question that reading tenure letters is often something of a hermeneutical exercise. It might also be the case that the interpretive work required to reveal their secrets would be eased

if people were willing to be more directly negative or if more negatively inclined people were induced to write. Generally speaking, however, sandwiched between the opening “throat-clearing” paragraph and the concluding line that almost invariably recommends promotion, referees offer ample clues to their real feelings about the case. Does the writer critically engage with the candidate’s work and try to explain why it is important, or influential, or even wrong-headed (a good sign)? Or does the letter merely recite the contents of the candidate’s CV (a bad sign)? Is the tone enthusiastic (good) or dutiful (bad)? Is the candidate someone who was already familiar to the writer before the evaluation request (good), or does the letter open with something like, “I’d never heard of Professor X before, but based on the dossier you sent me, he seems to be pretty smart” (bad)? It is the faculty’s responsibility to read these letters with care and sensitivity to create a well-rounded picture of the candidate’s quality and standing in the field, not simply to tally votes and approve an appointment once a candidate’s file accumulates the requisite number of “ayes.” Paying honoraria for evaluations, as Weyland suggests, might generate a wider range of evaluations. Even then, however, positive letters likely would still come in a range of varieties, and the responsibility for careful reading and reasoned judgment would remain.

There are also other valuable sources of information in the letters. As Weyland notes, it is often possible to make inferences about the candidates from the patterns of acceptances and declinations of invitations to review. If one has written to the right people—that is, to external colleagues who are themselves leading figures in a candidate’s precise area of scholarship, where she or he can be reasonably expected to be a known figure—and many of them refuse to write, that alone is an indicator of the candidate’s standing. To be sure, as Weyland argues, this type of inference from silence might not be necessary if more people were induced to write—but the level of responsiveness from the subfield community can still be telling. Moreover, a case for which an institution must approach too many people in order to obtain the requisite number of letters to move forward, or for which leading figures in the field consistently decline to write, generally merits close scrutiny; something deeper is probably amiss.

Finally, useful letters often compare the candidate to others in the same field (i.e., sometimes referees are explicitly asked for comparisons, and some institutions provide a list of comparators). Many people find these comparisons off-putting, but they can be helpful in locating a candidate within an array of scholars working on similar topics to assess impact and professional standing. Tenure review is almost unique among familiar peer-review processes in academic life in that it focuses attention on a decision about a singular case, which leads departments and academic leaders to lean on something that appears to be an absolute standard (i.e., whether the candidate is “above the bar,” to use a common metaphor). However, other common peer-review processes—for example, evaluating grant applications or papers submitted to a journal—also involve relative judgments. Not only is this submission fundable (or publishable), but is it also among the best so that it merits the allocation of a scarce resource (e.g., money or pages)? We should think about tenure evaluations in the same way: Are the candidates among the best in their field? Again, even generally positive letters can provide useful guidance.

I am not suggesting that institutions should not consider paying honoraria for writing letters (as some already do). This approach might provide, as Weyland suggests, enough incentive

for more negative or on-the-fence referees to write and to express a broader portfolio of views. Inducing more honest (and, thus, presumably mixed) language from referees might render it easier for departments, promotion and tenure committees, and academic leaders to make difficult calls in borderline cases. However, given what is at stake—for both the candidate and the institution—tenure decisions merit more careful consideration than simply totaling up pro and con recommendations in letters. Ultimately, we should remember that external-review letters are a supplement to rather than a substitute for our own careful and critical judgment. ■

RESPONSE TO SPOTLIGHT ON PROMOTION LETTERS: AN OPTION WORTH EXPLORING

Cynthia Ophelm, *Texas State University*

DOI: 10.1017/S1049096518002172

The suggestion that external reviewers of tenure and promotion candidate files, in general, have become less discerning is an important issue for the discipline. My perspective on this question comes from being a senior administrator and serving several years as associate provost. Thus, my comments are more general and comparative in nature.

There are two questions to address: Is there a self-selection process that results in generally positive and less-useful external reviews? Would a significant honorarium for reviewers reverse or ameliorate this trend? There indeed may be self-selection toward more general and/or positive reviews. My discussions with deans and chairs over the years, as well as in deliberation of this proposal, lead me to suspect that reviews tend to be positive, that reviews for borderline candidates are more difficult to solicit, and that at least some reviews resemble recommendation letters rather than critical analyses of a promotion portfolio. That said, there often are significant nuances in external reviews that give pronounced clues. A borderline candidate who meets minimum standards often elicits a lower level of enthusiasm even though the review is positive.

To a certain extent, the tendency toward positive reviews is unavoidable. The primary culprit is the trend toward academic specialization. Faculty hired in tenure-track positions are encouraged to focus on a specific “body of work” and become part of a national network of scholars to attain maximum visibility and prestige. This necessarily means that at least some letters likely will be solicited from a select pool of colleagues reluctant to criticize someone they know or with whom they have worked. Another inevitable trend for tenure candidates, particularly in those universities striving for higher research profiles, is that these candidates are increasingly “weeded out” in the years before the tenure decision. Many chairs tell me they have fewer borderline tenure candidates than in previous years.

The suggestion that senior scholars may be hesitant to take on borderline cases because they fear liability issues cannot be dismissed. Confidentiality notwithstanding, many states—including Texas—interpret public-information statutes as allowing unsuccessful candidates access to external reviews. In theory, this may lead to defamation suits; in practice, lawsuits based on external comments are rare. However, if someone is concerned with legal liability, an honorarium is not likely to make a difference.