Letters of Reference

It is the time of year again (as of the writing of this column) for writing and reading letters of recommendation for undergraduates applying to graduate school, postdoctoral associates seeking to join new research groups, and graduates applying for faculty positions. As I am sure many of you have experienced, this can take a lot of time and effort, to the point that [1] argues that perhaps it is time to get rid of recommendation letters but I doubt that will happen any time soon.

Given the importance of these letters to future careers, it was disappointing to read the conclusions of a recent study of the gender differences in recommendation letters for postdoctoral fellowships in geosciences [2]. The authors examine the relationship between applicant gender and two important letter attributes: length and tone. The relevance of letter tone is obvious, but letter length is typically positively associated with letter quality (both the actual letter quality and, I think, how the letter quality is perceived by the reader). The authors of [2] define a "long letter" as being at least 50 lines long. The sample size of the letter pool was large (1224 letters), with authors from 54 different countries, making this one of the largest known studies of gender bias in recommendation letters in any STEM field. As such, the results have implications outside of geosciences and, thus, are likely to be highly relevant in engineering as well. The analysis suggests "that women are significantly less likely to receive excellent recommendation letters than their male counterparts at a critical juncture in their career" [2].

Digital Object Identifier 10.1109/MCS.2016.2643218
Date of publication: 17 March 2017

A well-written letter should provide a clear indication for why the claims are being made.

Gender biases have been well documented in the fields of science and technology [3], [4], but [2] doesn't provide much discussion of the likely reasons for the biases in these letters, other than to note that recommendation letters offer a personal assessment, so it is possible that "biases of the writer are more likely to surface." These implicit biases can arise through being women described as "less confident and forceful," while being "more nurturing than men" [2]. Biases also arise through the use of fewer *standout adjectives* (for example, women are more likely to be described as "hardworking" and/or "diligent" versus "superb" and "brilliant") [5], [6].

The discussion in [7] suggests that there are three steps to help prevent biases from influencing behaviors when writing and evaluating letters: "You have to be aware that you have the biases in the first place, you have to be motivated to set them aside, and you have to have the time and effort to do so." I hope that each IEEE Control Systems Society member involved in the assessment process will pay careful attention to this advice in future letter writing and evaluations.

WHAT MAKES AN EXCELLENT LETTER?

The "methods" section of [2] provides good insight on the differences between letters that were evaluated as having a tone that was "excellent" or simply "good." In particular, excellent letters use numerous standout adjectives such as brilliant, rising star, pioneer, genius; praise the candidate's potential to become a leader in the field; and indicate ways that the candidate is superior to others. Excellent letters can include some elements of doubt, but these are outweighed by the positives. Good letters, on the other hand, while generally positive and referring to the candidate's characteristics of being solid, knowledgeable, and able to learn new topics, typically do not provide strong comparisons with others candidates in the field or comment positively about leadership potential.

Thus, it is possible that a letter writer thinks the letter describes the candidate as being "excellent," while it is read as being only a "good" recommendation. A further issue I often have when evaluating letters is *calibration*. When drawing upon a large, international candidate pool, it is often the case that the letter writers are not well known to me, so how do I know that we are similarly calibrated as to what makes an excellent candidate? To address this, I look for an honest assessment with enough evidence to support the claims.

A well-written letter should provide a clear indication for why the claims are being made. For example, if the candidate's contribution is "unique," what is really unique about it? If the approach

For faculty applications and promotion letters, I try to include this detailed evidence by reading and providing my analysis of the material.

is thought to be "innovative," what is the particular innovation? The ultimate goal is to give readers sufficient evidence and detail so that they can reach the same level of confidence in the assessment as the author, rather than just accepting the letter writer's recommendation at face value.

The same holds for undergraduate student letters of recommendation. I often see superlatives used to describe students with insufficient support to justify the statements; this leaves me wondering—is it just hyperbole? Establishing and comparing the candidate to his or her cohorts makes the recommendations and assessment particularly helpful. As a colleague once pointed out, "even at MIT, half of the students are below average."

With student applications, I try to provide anecdotes from interactions that help illustrate the specific research and technical skills that are of particular interest. For faculty applications and promotion letters, I try to include this detailed evidence by reading and providing my analysis of the material. A key issue with letters for faculty applications or promotions is assessing the impact (or at least the potential of it) of the work, often leading to the use of bibliometric indices.

BIBLIOMETRIC INDICES

However, the problems with using these indices are well documented [8], so they must be interpreted very carefully. The issue is that the numerical values are not necessarily good indicators of what is actually being assessed, which is the *impact* of the work. However, impact is difficult to determine, and the result typically varies from one observer to another. For academics in your own field, it is typically easy to assess this by determining the extent to which the work has impacted your own research. Outside of my field, I typically try to assess the extent to which other research groups have followed and cited the work, but that can be difficult. A great letter writer would dig deep in the literature to identify how the work in question has actually impacted the field—either as a transition (or "graduation" as I am now hearing it called) leading to a deployment in industry or by forming the foundation for further academic work. These types of impacts should lead to numerous citations if influential, but, as noted previously, these citations are counted just as much in the numerical scores as the brief mentions.

These assessments are important but can take a lot of time and effort—typically more than a full day, in my experience, to write a good promotion letter. The good news is that new tools are being developed to help assess academic work. ResearchGate [9] provides a mechanism to share research, and it also tracks metrics such as "reads," which gives a more contemporary perspective than the standard Google Scholar citation count. It is possible to interpret reads as a measure of other researchers' interest in the work, but this can have similar flaws to the citation indices.

One of the newest tools is Semantic Scholar [10], [11], with its tag line of "cutting through the clutter." Semantic Scholar was designed to try to measure the *influence* that a scientist or organization has on subsequent research. Researchers are ranked by an influence score, which provides a measure of the work's impact on future research. It probably comes as no surprise that the technique is based on machine-learning tools that were used to teach an algorithm how to read the different sections of a paper. The algorithm then analyzes several factors, including the number of citations to a publication and the surrounding context for these citations.

The result is an influence graph that depicts authors who have influenced the person and who that person has influenced. It also gives an indication of the number of highly influential citations that person has published. It has yet to be seen if this new tool will become an accepted indicator of impact. Even though the database size is limited, it already appears to be a very useful addition to the assessments tool already available. I recommend that you test it out next time you are writing a recommendation letter and see if its measures of impact are consistent with your own perspectives.

As always, I look forward to reading your feedback and insights.

REFERENCES

[1] R. Wikes, "Retire the letter of reference," *Science*, vol. 351, no. 6273, p. 630, Feb. 5, 2016. doi: 10.1126/science.351.6273.630.

[2] K. Dutt, D. L. Pfaff, A. F. Bernstein, J. S. Dillard, and C. Block, "Gender differences in recommendation letters for postdoctoral fellowships in geoscience," *Nature Geosci.*, vol. 9, pp. 805–809, Nov. 2016. doi:10.1038/ngeo2819.

[3] V. Valian, *Why So Slow? The Advancement of Women*. Cambridge, MA: MIT Press, Dec. 1997.

[4] C. Hill, C. Corbett, A. St. Rose. (2010, Feb.). Why so few? Women in science, technology, engineering, and mathematics. American Association of University Women (AAUW). [Online]. Available: http://www.aauw.org/files/2013/02/Why-So-Few-Women-in-Science-Technology-Engineering-and-Mathematics.pdf

[5] F. Trix and C. Psenka, "Exploring the color of glass: Letters of recommendation for female and male medical faculty," *Discourse Soc.*, vol. 14, no. 2, pp. 191–220, 2003.

[6] T. Schmader, J. Whitehead, and V. Wysocki, "A linguistic comparison of letters of recommendation for male and female chemistry and biochemistry job applicants," *Sex Roles*, vol. 57, no. 7/8, pp. 509–514, 2007.

[7] R. Skibba. (2016. Oct. 3). Women postdocs less likely than men to get a glowing reference. *Nature*. [Online] Available: http://www.nature.com/news/women-postdocs-less-likely-than-men-to-get-a-glowing-reference-1.20715

[8] R. D. Braatz, "Chasing impact factors or making an impact on technology?" *IEEE Control Syst. Mag.*, vol. 32, no. 6, pp. 6–7, Dec. 2012.

[9] [Online] Available: researchgate.net

[10] J. Bohannon, "Who's the Michael Jordan of computer science? New tool ranks researchers' influence," *Science*, Apr. 2016. [Online]. Available: doi: 10.1126/science.aaf9939

[11] [Online] Available: https://www.semanticscholar.org/

Jonathan P. How