Help That Hurts Women

Some scholars have questioned academe's reliance on letters of recommendation, saying they're <u>onerous for the professors writing them</u> or <u>speak more about connections</u> to "big-name" scholars than substance, or both.

A recent study explores another concern about letters of recommendation: whether they're biased against the women they're supposed to help. The short answer is yes.

The longer answer -- and the study's obvious takeaway for recommendation-letter writers and readers -- is that letters about women include more doubt-raising phrases than those about men, and that even one such phrase can make a difference in a job search.

"Doubt raisers," as the study calls them, come in different forms. But all suggest less than complete confidence in the letter seeker. First, there are "hedges," as in, "She might be good." Then there are "negatives," such as, "She's not great," along with "faint praises," such as, "She'll do OK." Last, irrelevant information also raises red flags for a reader.

Both male and female letter writers are guilty of this largely unconscious bias, the new paper says. However unintended, the consequences are real. Faculty members reading the letters included in the study noticed the presence of even one doubt raiser and evaluated subjects of the letters more negatively.

These results build on other research about gendered language in recommendation letters, including a 2016 study that found letter writers use language that portrays female academic job applicants as less dynamic and excellent than their male counterparts.

Shortchanging Women

"People should be aware that they may be shortchanging women by inadvertently using doubt raisers in their letters of recommendations for them," said Michelle R. Hebl, Martha and Henry Malcolm Lovett Chair of Psychology at Rice University, a co-author on the new study. "In essence, writers may be more likely to describe women and men as, 'She has the potential to be good,' whereas, 'He is good,' or, 'She may not be the best leader but she is competent,' whereas, 'He is competent.'"

For their two-part study, Hebl and her colleagues analyzed 624 letters of recommendation for 174 job applicants to eight junior faculty positions in psychology at an unnamed research-intensive university in the South. About half the applicants were male and half were female. Approximately 30 percent of recommenders were female and about 70 percent were male. The applicants' mean age was 32 and the mean number of letters per job seeker was 3.6.

Research coders who were, by design, "blind" to the writers' and applicants' genders and other identifying information, scoured each letter for the four different kinds of doubt raisers and rated them accordingly. To make sure that men didn't simply inspire more confidence in their letter writers, researchers also controlled for 10 academic performance variables found on applicants' CVs, such as number of first-authored papers, number of honors, number of years spent as a postdoctoral fellow, positions applied for and their graduate institutions' national rankings.

An advanced analysis found no difference by gender in these performance variables. But the study's authors were able to predict the content of the letters -- meaning the presence and frequency of doubt raisers -- by gender. Over all, letters for female applicants had an average of 0.69 doubt raisers and those for men had 0.55.

Across genders, 52 percent of the letters had at least one doubt raiser, 10 percent had two or more and 48 percent of the letters had no doubt raisers. But among female applicants, 54 percent of letters had at least one, 13 percent had two or more and 46 percent had none. Among men's letters, 51 percent had at least one, 7 percent had two or more, and 49 percent had none.

The most common kind of doubt raiser across letters was faint praise (27 percent of the sample had at least one instance), followed by hedging (18 percent of the sample contained one or more), irrelevancy (14 percent) and negativity (12 percent).

For female applicants, 30 percent of letters had faint praise, 20 percent had hedging, 14 percent had negativity and 12 percent had irrelevancy. Among male applicants, 24 percent of letters had faint praise, 16 percent had irrelevancy, 15 percent had hedging and 10 percent

had negativity. The authors say the relationship between applicant gender and doubt raisers is significant for all types of phrases except irrelevant information.

For both men and women, most doubt raisers were about research productivity, as opposed to teaching.

Hidden Signals

Next, the study's authors sought to test whether doubt raisers caused readers of recommendation letters to evaluate applicants more negatively.

Hebl and her team recruited 305 professors from across the U.S. to participate in an online study. The majority were psychologists in terms of field and full professors with regard to rank. Participants were told they'd be reading a redacted recommendation letter. But they didn't know that embedded in the four-paragraph letter was a doubt raiser such as, "I can say with certainty that 'AA' does not have the skills to be the best researcher you have ever seen, but she/he does have the potential to become successful in developing an independent research program at your institution." Or, "I have confidence that 'AA' will become better than average at being successful in developing an independent research program at your institution."

The faculty readers eventually evaluated the hypothetical applicants for research and teaching potential. As the study's authors expected, the letters with doubt raisers were evaluated more negatively with regard to research potential. The effects were the same for both male and female applicants.

"Doubt raisers are a minus for everyone, but letter writers assign that minus more often to women than to men," the study says. "If search committees ignored letters of recommendation, that asymmetry would not matter. But letters of recommendation are commonly used as selection tools in academia." And the data "have important implications for women in academia, particularly because women face biases early in the selection process."

The inclusion of even a single doubt raiser -- particularly negativity or hedging -- "was enough to lead to statistically lower evaluations of the applicant," the study says. That's salient because, again, the first part of the study showed that 14 percent and 20 percent of the letters for female applicants had at least one negativity or hedging doubt raiser, respectively, compared to 10 percent and 15 percent of the letters for the male applicants.

"Although these gender differences, while reliable, are small," the study concludes, they show that "only one statement can make a difference for an applicant."

Hebl said she hoped the study makes letter writers "more aware" of their word choices. They may also do well to "proofread their letters so that they write letters that are just as strong for women as they are for men."

Hebl's co-authors were Juan M. Madera, an associate professor of management at the University of Houston; Heather Dial, a postdoctoral fellow in communication sciences at the University of Texas at Austin; Randi Martin, Elma Schneider Professor in the department of

psychology at Rice; and Virginia Valian, distinguished professor of psychology at Hunter College of the City University of New York. The work appears in the *Journal of Business and Psychology*.