



# Gender Bias Interacts With Instructor Disfluency to Negatively Affect Student Evaluations of Teaching

Jessica LaPaglia\*, Katelyn Miller and Samantha Protexter

Department of Social Sciences, Morningside University, Sioux City, IA, United States

## OPEN ACCESS

### Edited by:

David Gonzalez-Gomez,  
University of Extremadura, Spain

### Reviewed by:

Judit Bóna,  
Eötvös Loránd University, Hungary  
Ehsan Namaziandost,  
Mehrarvand Institute of Technology,  
Iran

### \*Correspondence:

Jessica LaPaglia  
lapagliaj@morningside.edu

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Education

**Received:** 17 November 2021

**Accepted:** 24 January 2022

**Published:** 24 February 2022

### Citation:

LaPaglia J, Miller K and  
Protexter S (2022) Gender Bias  
Interacts With Instructor Disfluency  
to Negatively Affect Student  
Evaluations of Teaching.  
*Front. Educ.* 7:817291.  
doi: 10.3389/feduc.2022.817291

Recent research has shown that instructor fluency can impact student judgments of learning and instructor ratings but has no real effect on actual learning. In addition, women tend to receive lower course and instructor evaluations than men. In the current study, we examined how instructor fluency and instructor gender influenced instructor evaluations and student learning. Participants watched a short lecture video. The speaker was either male or female and was either fluent (i.e., even paced in their speech) or disfluent (i.e., disorganized, made mistakes). Following the video, participants evaluated the instructor and took a quiz over the lecture. Results indicated that disfluency negatively affected quiz scores, but instructor gender did not. Participants rated the female speaker significantly lower than the male speaker, but only when the speaker was disfluent. These results are explained through the lens of attributional gender bias.

**Keywords:** instructor fluency, gender bias, student evaluations of teaching, learning, judgments of learning

## INTRODUCTION

Instructor and course ratings are often used in higher education to make decisions regarding promotion and tenure. However, research indicates that these evaluations can be biased against women and people of color (Chavez and Mitchell, 2020; Finn, 2020). In addition to the gender of the instructor, the fluency of the instruction can alter student perceptions of learning and instructor ratings (Carpenter et al., 2016). In the current study, we examined the potential interactions between instructor gender and the lecture delivery on student learning and instructor ratings.

The fluency of an instructor's lecture has been shown to influence students' judgments of learning (JOL), but not actual learning (Carpenter et al., 2013; Toftness et al., 2018). Carpenter and colleagues showed participants a video of a short lecture. The female instructor presented the lecture in a fluent manner (i.e., did not use her notes, maintained eye-contact, and stood up straight) or in a disfluent manner (used notes, looked away frequently, and slouched). They discovered that participants believed that they did not learn as much in the disfluent condition compared to the fluent condition, but their actual learning was equivalent across conditions. These findings were replicated by Carpenter et al. (2016) who also found that fluency could affect student perceptions of instructors in addition to their perceptions of learning. Specifically, students rated the fluent instructor significantly higher on organization, knowledge, preparedness, and effectiveness compared to the disfluent instructor. Recently, this finding has not been shown

to be moderated by the instructor's apparent experience (Carpenter et al., 2020a). Carpenter and colleagues have primarily manipulated visual forms of disfluency (e.g., poor eye-contact and slouching). However, differential effects of fluency have been noted in the literature. Some types of disfluency, such as perceptual disfluency involving cursive font types, can improve learning (Geller et al., 2018). In the present experiment, we used auditory forms of disfluency which has yet to be examined within the context of instructor ratings.

The findings of previous research (e.g., Carpenter et al., 2020a) leave two questions left unanswered. The first is whether verbal, as opposed to visual, forms of disfluency influence student perceptions of learning and the instructor. The second is whether the effects of instructor fluency are moderated by the gender of the instructor. In the studies by Carpenter et al., the disfluent conditions are disfluent because the instructor slouches, flips through notes, and has poor eye contact with the audience. Although the instructor also has more halted speech, the primary modes of disfluency are visual. In the current study, we sought to examine how solely verbal/auditory forms of disfluency might affect student learning, perceptions of learning, and instructor ratings. This verbal disfluency involved the instructor using a lot of "ums," accidentally skipping and going back to slides, and speaking quickly at times in the lecture. While speaking quickly can be an indication of fluency, in this case it altered the pace of the lecture. Although the information presented in the fluent and disfluent conditions is the same, there is reason to believe that verbal disfluency could negatively influence learning. For instance, when text is coherent, it tends to be more easily remembered than incoherent text (e.g., Rawson and Dunlosky, 2002). Moreover, non-native accented speech reduces listening comprehension compared to native accented speech (e.g., Major et al., 2002). Therefore, information that is difficult to process will likely lead to decreases in both perceived and actual learning.

Carpenter and colleagues have shown that fluency influences student perceptions of learning both when the instructor was a male (Toftness et al., 2018) and when the instructor was a female (Carpenter et al., 2013), but this has yet to be manipulated within the same experiment. Research on gender bias in education is vast and indicates a bias against women in student evaluations of teaching (Basow and Silberg, 1987; Martin, 2016; Boring, 2017; Rosen, 2017; Mitchell and Martin, 2018; Mengel et al., 2019). This gender bias has even been shown in highly controlled studies. For instance, MacNell et al. (2015) found that in an online class where participants were told that their instructor was either a male or a female, the female identity was rated significantly lower than the male identity regardless of the actual gender of the instructor. It is unclear whether fluency would affect ratings of instructors differently depending on their gender.

In the present experiment, participants watched a lecture video that was voiced by either a male or a female instructor. This instructor was either well-spoken (fluent condition) or disorganized (disfluent condition). Participants made a JOL and rated the instructor on a variety of measures (e.g., knowledge of subject matter). Following this survey, they took a quiz on the content of the lecture. Given the extensive research on gender bias, we hypothesized that the female instructor will be rated

lower than the male instructor in both the fluent and disfluent conditions. Consistent with the research by Carpenter et al. (2013, 2016, 2020a), we further hypothesized that participants in the fluent condition would overestimate their learning and rate the instructor higher compared to the disfluent condition. Fluency could have no effect on quiz performance (consistent with work done in the Carpenter lab); however, because the disfluency of the current study involves more verbal, as opposed to visual, disfluency, we may find that participants perform worse on the quiz in the disfluent condition compared to the fluent condition.

## MATERIALS AND METHODS

### Participants

There were 72 participants (49 female, 23 male) from a small Midwestern university who participated in this experiment for course credit. Participants signed up to participate in this study via the psychology department's research participation webpage. They were primarily students from the general psychology course that was made up of freshmen and sophomore students from all majors. Their mean age was 19.40 ( $SD = 1.67$ ). There were 18 participants in each condition (female-fluent, female-disfluent, male-fluent, and male-disfluent). Participant gender ratios were nearly identical in each condition.

### Materials and Procedure

After obtaining informed consent, the experimenter presented participants with a video about the production of cocoa. Participants watched the video while sitting in a cubical with headphones to minimize distractions. The video lecture was a PowerPoint presentation with images and minimal text. The experimenter instructed participants to pay attention to the lecture video because they would receive a quiz over the video later. Participants were randomly assigned to watch one of four versions of the lecture video. Two of the lectures were voiced by a male and two by a female. The male and female each recorded two videos, one was fluent and the other disfluent. In the fluent condition, the speaker read the lecture script at an even pace and spoke clearly. In the disfluent condition, the speaker sometimes spoke quickly, frequently skipped and went back to slides, coughed into the microphone, and generally sounded disorganized. Here, disfluency is defined as phenomena that are not typical in speech. See **Table 1** for an excerpt of the fluent and disfluent scripts. The speaker was never visible in the video. The videos were approximately 6 min each.

Following the video, participants completed a survey in Google Forms. In the survey, participants made a JOL in which they guessed the percentage correct that they would get on a quiz about the lecture momentarily. They also rated the instructor's organization and knowledge of the topic, their interest in the topic, motivation to learn the information, and provided an overall rating of the instructor and the lecture. Participants were also asked whether they recognized the voice in the video. No participant had recognized the instructor's voice. A quiz immediately followed the survey. The quiz consisted of 10, four-option multiple choice questions. These questions probed factual

**TABLE 1** | Excerpts from the video scripts.

Fluent condition	Disfluent condition
<p>“Cocoa beans grow in football-sized pods that sprout off of the trunk and branches of cocoa trees. Initially green in color, the pods turn yellow-orange when they are ripe. Workers harvest the pods by hand, using machetes and long poles with cutting edges for the harder-to-reach pods. Each pod contains about 30–50 beans. Once the cocoa pods have been collected in baskets, they are transported to a processing house where they are cut open and the cocoa beans are removed.”</p>	<p>“Um, oh yeah, cocoa beans grow in football-sized pods... wait, is it football sized? Yeah. Cocoa beans grow in football-sized pods that sprout off of the trunk and branches of cocoa trees (read next sentence quickly). Initially green in color, the pods turn yellow-orange when they are ripe. Workers harvest the pods by hand, using machetes and long poles with cutting edges for the harder-to-reach pods. Each pod contains uh, about 30–50 beans (pause, pretend to flip through a book). Once the cocoa pods have been collected in baskets, they are transported to a processing house where they are cut open and the cocoa beans are removed.”</p>

information presented in the lecture video (e.g., “Which type of cocoa bean tree supplies the majority of the world’s chocolate?”). The entire experiment took approximately 10 min.

## RESULTS

### Metacognition and Learning

A 2 (instructor gender: male, female)  $\times$  2 (fluency: fluent, disfluent) between subjects ANOVA revealed no significant interactions or main effects for JOLs,  $ps > 0.394$ . Overall, participants were highly confident with a mean predicted quiz score of 71.8%. The actual mean quiz score was 69.4%, but there was no significant correlation between predicted and actual quiz score,  $r = 0.139$ ,  $p = 0.244$ . Although the predicted proportion correct in the fluent condition was slightly higher ( $M = 0.74$ ,  $SD = 0.16$ ) than in the disfluent condition ( $M = 0.70$ ,  $SD = 0.19$ ), this difference was not significant,  $p = 0.390$ , likely due to low statistical power (power = 0.53).

Although JOLs were not affected by either variable, there was a significant main effect of fluency when examining the dependent variable of number correct on the final quiz,  $F(1, 68) = 4.37$ ,  $p = 0.040$ ,  $\eta^2_p = 0.06$ . Participants scored better on the quiz when the lecture was delivered fluently ( $M = 7.42$ ,  $SD = 1.59$ ) than when it was delivered disfluently ( $M = 6.47$ ,  $SD = 2.17$ ). No other main effects or interactions were significant for quiz scores,  $ps > 0.33$ .

### Student Evaluations of Teaching

There was a significant interaction between speaker gender and fluency for instructor rating,  $F(1, 68) = 4.62$ ,  $p = 0.035$ ,  $\eta^2_p = 0.064$ . Means and standard deviations are presented in **Table 2**. When the instructor presented the lecture fluently, there was no difference in ratings between the male and female instructors. However, when the lecture was disfluent, participants rated the male instructor higher than the female instructor. There were also significant main effects of instructor gender,  $F(1, 68) = 4.62$ ,  $p = 0.035$ ,  $\eta^2_p = 0.064$ , and fluency,  $F(1, 68) = 98.82$ ,  $p < 0.001$ ,  $\eta^2_p = 0.592$ , with the male instructor rated higher ( $M = 3.31$ ,  $SD = 1.19$ ) than the female instructor ( $M = 2.86$ ,  $SD = 1.53$ ) and the fluent instructor rated higher ( $M = 4.11$ ,  $SD = 0.82$ ) than the disfluent instructor ( $M = 2.06$ ,  $SD = 1.01$ ).

For lecture rating, there was no significant interaction,  $p = 0.207$ . However, there was a significant main effect of lecture fluency with the fluent lecture ratings higher ( $M = 4.00$ ,  $SD = 0.79$ ) than the ratings of the disfluent lecture ( $M = 2.22$ ,

$SD = 1.07$ ),  $F(1, 68) = 66.57$ ,  $p < 0.001$ ,  $\eta^2_p = 0.495$ . There was also a marginally significant main effect of instructor gender with the male instructor’s lecture receiving slightly higher ratings ( $M = 3.31$ ,  $SD = 1.14$ ) than the female instructor’s lecture ( $M = 2.92$ ,  $SD = 1.42$ ),  $F(1, 68) = 3.19$ ,  $p = 0.079$ ,  $\eta^2_p = 0.045$ .

In examining student ratings of the instructor’s level of organization, there was a significant interaction,  $F(1, 68) = 6.39$ ,  $p = 0.014$ ,  $\eta^2_p = 0.086$ . When the lecture was presented fluently, the female instructor was rated as more organized than male instructor. When the lecture was presented disfluently, there was a reversal and the male instructor was rated as more organized than the female instructor. The main effect of instructor gender was not significant ( $p = 0.265$ ), but not surprisingly, the instructor was rated as more organized in the fluent condition ( $M = 4.47$ ,  $SD = 0.65$ ) than in the disfluent condition ( $M = 2.14$ ,  $SD = 1.05$ ),  $F(1, 68) = 139.16$ ,  $p < 0.001$ ,  $\eta^2_p = 0.672$ . For ratings of instructor knowledge, there was no significant interaction or main effect of gender,  $ps > 0.125$ . However, the disfluent instructor was rated significantly less knowledgeable ( $M = 2.39$ ,  $SD = 1.27$ ) compared to the fluent instructor ( $M = 4.58$ ,  $SD = 0.60$ ),  $F(1, 68) = 88.93$ ,  $p < 0.001$ ,  $\eta^2_p = 0.567$ .

Participants rated their interest in the topic and motivation to learn the material. For interest, there was no significant interaction or main effect of instructor gender,  $ps > 0.100$ . However, participants were significantly more interested in the material when it was presented fluently ( $M = 3.50$ ,  $SD = 1.03$ ) compared to when it was presented disfluently ( $M = 2.75$ ,  $SD = 1.11$ ),  $F(1, 68) = 8.99$ ,  $p = 0.004$ ,  $\eta^2_p = 0.117$ . For motivation

**TABLE 2** | Means (and standard deviations) for ratings of instructor, lecture, organization of the instructor, knowledge of the instructor, interest in the lecture, and motivation to learn the information as a function of instructor gender and fluency.

	Fluent		Disfluent	
	Female	Male	Female	Male
Instructor	<b>4.11 (0.90)</b>	<b>4.11 (0.76)</b>	<b>1.61 (0.85)</b>	<b>2.44 (1.04)</b>
Lecture	3.94 (0.87)	4.06 (0.73)	1.89 (1.08)	2.50 (1.04)
Organization	<b>4.61 (0.50)</b>	<b>4.33 (0.77)</b>	<b>1.78 (0.88)</b>	<b>2.50 (1.10)</b>
Knowledge	4.67 (0.59)	4.50 (0.62)	2.11 (1.23)	2.67 (1.28)
Interest	3.28 (0.96)	3.72 (1.07)	2.56 (0.98)	2.94 (1.21)
Motivation	3.00 (0.97)	3.39 (0.98)	2.28 (1.13)	2.61 (1.09)

*Significant ( $p < 0.05$ ) interactions are bolded.*

to learn the material, there was no significant interaction or main effect of instructor gender,  $ps = 0.147$ . Participants who viewed the fluent lecture were more motivated to learn the material ( $M = 3.19$ ,  $SD = 0.98$ ) compared to those who viewed the disfluent lecture ( $M = 2.44$ ,  $SD = 1.11$ ),  $F(1, 68) = 9.28$ ,  $p = 0.003$ ,  $\eta^2_p = 0.120$ . Participant gender did not influence any dependent variables measured in this study.

## DISCUSSION

In the present experiment, we examined the influence of instructor gender and fluency on instructor ratings and student learning. Contrary to previous research (e.g., Carpenter et al., 2013), we found that fluency had no effect on JOLs. Participants were highly confident in their learning regardless of instructor fluency or gender. Another finding divergent from the work of Carpenter and colleagues was that disfluency reduced quiz performance. Consistent with our hypothesis, we found a gender bias in instructor ratings, but only when the instructor was disfluent. We break down these key findings in the following sections.

### Metacognition and Learning

Carpenter et al. (2013) found that instructor disfluency reduces JOLs but has no effect on actual learning. In the present experiment, disfluency had no effect on JOLs, but reduced learning compared to the fluent condition. What, then, can account for these divergent findings? The answer lies in the type of disfluency used. Carpenter and colleagues used primarily visual-based disfluency (e.g., slouching, poor eye-contact) while we used auditory-based disfluency (e.g., “ums,” sounding unsure, speaking too quickly). The auditory disfluency used in the present study disrupted the message, thus reduced learning. Indeed, when instructors lack clarity in their message, it can increase cognitive load in students and negatively affect learning (Bolkan, 2016).

In terms of JOLs, we did not replicate Carpenter et al. (2013) who showed that fluent instruction lead to overconfidence. It appears that the current sample, although slightly overconfident, had a mean JOL that was close to the mean quiz score regardless of condition—perhaps due to the type of quiz. Carpenter et al. (2013) used a free-recall final test whereas we used a multiple-choice quiz. Free recall tends to be more challenging because the correct answer needs to be retrieved, but only recognition of the correct response is necessary in multiple choice (Kintsch, 1970). Carpenter et al. (2020a) also used multiple choice, but their quiz consisted of 30 items as opposed to 10 in the present experiment. Indeed, performance in their study was around 55% whereas our participants neared 70% correct on the final quiz. Therefore, we are likely seeing ceiling effects in performance, thus reducing the ability to detect differences between the conditions.

### Student Evaluations of Teaching

A gender bias in student evaluations of teaching was found only in the disfluent condition. Research in attributional gender bias provides insight into why we might expect instructor ratings to vary for men and women instructors in the different fluency

conditions. People make attributions about behavior differently for men and women. For instance, Espinoza et al. (2014) found that teachers attribute poor math performance in girls as a lack of ability (internal forces) whereas poor math performance in boys is more likely to be attributed to external factors, such as a lack of effort. Likewise, women in some leadership positions receive less internal and more external attributions for their success than men (Lopez and Ensari, 2014). In the present experiment, we suspect that when the instructor was disfluent, there was a gender bias because participants attributed the female instructor’s disfluency to internal forces (i.e., a lack of ability, knowledge) whereas the male instructor’s disfluency was attributed to external forces (i.e., not enough time to prepare the lecture). Thus, the female instructor performed poorly because she is a bad teacher and the male instructor performed poorly because he was having a bad day. It is important to note that in a typical classroom situation, the students would see many lectures from their instructor and whether the instructor is typically a strong teacher or not should be more apparent.

## CONCLUSION

Student evaluations of teaching are widely used as a measure of teaching effectiveness despite being easily influenced by other factors such as gender of the instructor and grade expectations (Boring et al., 2016; Finn, 2020). Students’ ideas of what helps them learn are often incorrect (Finn and Tauber, 2015). If there is an over-reliance of student evaluations of teaching in promotion and tenure decisions, it could lead to instructors using methods that will increase these subjective ratings rather than enhance student learning (Carpenter et al., 2020b; Oppenheimer and Hargis, 2020). The current study extended research from Carpenter et al. (2020a) to clarify that disfluency can affect instructor ratings and highlighted that female instructors may be rated much lower than male instructors when they are ill-prepared for a lecture even though learning suffers in a disfluent lecture regardless of the gender of the instructor. One silver lining here was that when the lecture was presented fluently, there was no gender bias, suggesting that gender biases are more prominent when learning is challenged by instructor disfluency. This research provides further evidence that female instructors are disadvantaged by student evaluations of teaching and highlights the need for other forms of evaluating teaching. Formative assessment, evaluations of teaching portfolios, and peer teaching observations may provide useful feedback to instructors that will lead to changes instruction to increase actual student learning rather than just student perceptions of learning (Gurung, 2020).

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found online through the Open Science Framework: [osf.io/r5mhj](https://osf.io/r5mhj).



## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Morningside University Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

## REFERENCES

- Basow, S. A., and Silberg, N. T. (1987). Student evaluations of college professors: are female and male professors rated differently?. *J. Educ. Psychol.* 79, 308–314.
- Bolkan, S. (2016). The importance of instructor clarity and its effect on student learning: facilitating elaboration by reducing cognitive load. *Commun. Rep.* 29, 152–162. doi: 10.1080/08934215.2015.1067708
- Boring, A., Ottoboni, K., and Stark, P. B. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Res.* 1–11. doi: 10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1
- Boring, S. K. (2017). Gender biases in student evaluations of teaching. *J. Public Econ.* 145, 27–41. doi: 10.1016/j.jpubeco.2016.11.006
- Carpenter, S. K., Mickes, L., Rahman, S., and Fernandez, C. (2016). The effect of instructor fluency on students' perceptions of instructors, confidence in learning, and actual learning. *J. Exp. Psychol. Appl.* 22, 161–172.
- Carpenter, S. K., Northern, P. E., Tauber, S. U., and Toftness, A. R. (2020a). Effects of lecture fluency and instructor experience on students' judgments of learning, test scores, and evaluations of instructors. *J. Exp. Psychol. Appl.* 26, 26–39. doi: 10.1037/xap0000234
- Carpenter, S. K., Witherby, A. E., and Tauber, S. K. (2020b). On students' (mis)judgments of learning and teaching effectiveness. *J. Appl. Res. Mem. Cogn.* 9, 137–151. doi: 10.1016/j.jarmac.2019.12.009
- Carpenter, S. K., Wilford, M. M., Kornell, N., and Mullaney, K. M. (2013). Appearances can be deceiving: instructor fluency increases perceptions of learning without increasing actual learning. *Psychon. Bull. Rev.* 20, 1350–1356. doi: 10.3758/s13423-013-0442-z
- Chavez, K., and Mitchell, K. M. W. (2020). Exploring bias in student evaluations: gender, race, and ethnicity. *Polit. Sci. Polit.* 53, 270–274. doi: 10.1017/S1049096519001744
- Espinoza, P., Arêas da Luz Fontes, A. B., and Arms-Chavez, C. J. (2014). Attributional gender bias: teachers' ability and effort explanations for students' math performance. *Soc. Psychol. Educ.* 17, 105–126. doi: 10.1007/s11218-013-9226-6
- Finn, B. (2020). What more can we learn from teaching evaluations?. *J. Appl. Res. Mem. Cogn.* 9, 157–160. doi: 10.1016/j.jarmac.2020.02.002
- Finn, B., and Tauber, S. K. (2015). When confidence is not a signal of knowing: how students' experiences and beliefs about processing fluency can lead to miscalibrated confidence. *Educ. Psychol. Rev.* 27, 567–586. doi: 10.1007/s10648-015-9313-7
- Geller, J., Still, M., Dark, V. J., and Carpenter, S. K. (2018). Would disfluency by any other name still be disfluent? Examining the disfluency effect with cursive handwriting. *Mem. Cogn.* 46, 1109–1126. doi: 10.3758/s13421-018-0824-6
- Gurung, R. A. R. (2020). Call it out: recognizing good teaching and learning. *J. Appl. Res. Mem. Cogn.* 9, 161–164. doi: 10.1016/j.jarmac.2020.02.003
- Kintsch, W. (1970). "Models for free recall and recognition," in *Models of Human Memory*, ed. D. Norman (New York: Academic Press), 333–373.

## AUTHOR CONTRIBUTIONS

JL conceptualized and designed the study, performed statistical analyses, and wrote the manuscript. KM and SP created the experiment materials and collected the data. All authors read and approved the submitted version of the manuscript.

- Lopez, E. S., and Ensari, N. (2014). The effects of leadership style, organizational outcome, and gender on attributional bias towards leaders. *J. Leadersh. Stud.* 8, 19–37. doi: 10.1002/jls.21326
- MacNell, L., Driscoll, A., and Hunt, A. N. (2015). What's in a name: exposing gender bias in student ratings of teaching. *Innov. High. Educ.* 40, 291–303. doi: 10.1007/s10755-014-9313-4
- Major, R. C., Fitzmaurice, S. F., Bunta, F., and Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: implications for ESL assessment. *TESOL Q.* 36, 173–190. doi: 10.2307/3588329
- Martin, L. (2016). Gender, teaching evaluations, and professional success in political science. *Polit. Sci. Polit.* 49, 313–319. doi: 10.1017/S1049096516000275
- Mengel, F., Sauermann, J., and Zolitz, U. (2019). Gender bias in teaching evaluations. *J. Eur. Econ. Assoc.* 17, 535–566. doi: 10.1093/jeea/jvx057
- Mitchell, K. M., and Martin, J. (2018). Gender bias in student evaluations. *Polit. Sci. Polit.* 51, 648–652. doi: 10.1017/S10490965180001X
- Oppenheimer, D. M., and Hargis, M. B. (2020). If teaching evaluations don't measure learning, what do they do?. *J. Appl. Res. Mem. Cogn.* 9, 170–174. doi: 10.1016/j.jarmac.2020.03.001
- Rawson, K. A., and Dunlosky, J. (2002). Are performance predictions for text based on ease of processing?. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 69–80. doi: 10.1037//0278-7393.28.1.69
- Rosen, A. S. (2017). Correlations, trends, and potential biases among publically accessible web-based student evaluations of teaching: a large-scale study of RateMyProfessors.com data. *Assess. Eval. High. Educ.* 43, 31–44. doi: 10.1080/02602938.2016.1276155
- Toftness, A. R., Carpenter, S. K., Geller, J., Lauber, S., Johnson, M., and Armstrong, P. I. (2018). Instructor fluency leads to higher confidence in learning, but not better learning. *Metacogn. Learn.* 13, 1–14. doi: 10.1007/s11409-017-9175-0

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 LaPaglia, Miller and Protexter. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.