Instructor Name Preference and Student Evaluations of Instruction

Melissa M. Foster, *The Ohio State University, USA*

ABSTRACT   Student evaluations of instruction (SEIs) have an important role in hiring, firing, and promotion decisions. However, evidence suggests that SEIs might be influenced by factors other than teaching skills. The author examined several nonteaching factors that may impact SEIs in two independent studies. Study 1 examined whether an instructor's name preference (i.e., first name versus "Dr." last name) influenced SEIs in actual courses. Study 2 implemented a two (i.e., instructor name preference: first name or "Dr." last name) by two (i.e., instructor gender: male or female) by two (i.e., instructor race: white or Black) between-subjects design for SEIs in a hypothetical course. Study 1 found that SEIs were higher when the female instructor expressed a preference for being called by her first name. Study 2 found the highest SEIs for Black male instructors when instructors asked students to call them by their first name, but there was a decrease in SEI scores if they went by their professional title. Administrators should be aware of the various factors that can influence how students evaluate instructors.

Understanding the values and limitations of student evaluations of instruction (SEIs) or student evaluations of teaching (SETs) is critical to the long-term success of higher education. SEIs have been used since at least the 1920s (Remmers and Brandenburg 1927) as a major criterion in hiring, firing, promotion, and tenure decisions (Seldin 1993). These evaluations also can indirectly impact an instructor's career by influencing factors such as course enrollment (Yu, Mincieli, and Zipser 2021).

However, the appropriateness of SEIs in measuring teaching effectiveness has been called into question for decades. Whereas a meta-analysis found a moderate-sized correlation ($r = 0.31$) between SETs and student achievement (Wright and Jenkins-Guarnieri 2012), SETs can be influenced by factors other than teaching quality (Henson and Scharfe 2011; Langbein 1994; Stroebe 2020; Youmans and Jee 2007). For example, correlations have been found between SEIs and the perceived easiness of a course, instructor physical attractiveness, course subject, and instructor gender (Rosen 2018). Additionally, SEIs are inconsistent and better at providing information about student evaluators

than the instructors (Clayson 2018). Moreover, there are measurement issues regarding which students are motivated to complete SEIs (Hoel and Dahl 2019), the likelihood of misleading conclusions due to acquiescence (Valencia 2019), small sample sizes (Holland 2019), and students not always responding honestly to SEI items (McClain, Gulbis, and Hays 2018). Even under the best-case scenario regarding bias and lack of reliability, SEIs alone are unlikely to provide high-quality information about instruction (Esarey and Valdes 2020)—and even after implementation of evidence-based teaching practices, SEIs may not change significantly (Stewart, Speldewinde, and Ford 2018). SEIs also may include inappropriate comments, which can harm the well-being of instructors (Lakeman et al. 2021).

One potential reason that nonteaching factors may influence SEIs is that there may be a relationship between these factors and perceived teacher friendliness and availability. The concept of *immediacy* refers to the degree of perceived physical or psychological closeness between people in a relationship (Mehrabian 1966, 1971). Applied to the classroom, teacher immediacy refers to the physical or psychological closeness between teachers and their students (Frymier 2013). Teachers can engage in various verbal behaviors (e.g., use personal examples, encourage students to ask questions, and refer to the class as "our" class) and nonverbal

Melissa M. Foster 🔾 *is senior lecturer of communication at The Ohio State University. She can be reached at foster.1182@osu.edu.*

behaviors (e.g., gesture while talking, move around the class, and maintain eye contact) to increase immediacy (Christophel 1990; Meyerberg and Legg 2015). One study found a significant correlation ($r = 0.54$) between ratings of teacher immediacy and teaching evaluations (Moore et al. 1996).

Some of the factors that influence perceptions of teacher immediacy are commonly considered skills for public speaking (e.g., use of personal examples and making eye contact), which could explain why there is a relationship between teacher immediacy and SEIs. However, other factors relevant to immediacy that are not associated with public-speaking skills could impact SEIs. Two of these factors are an instructor's name preference (i.e., being on a first-name basis with students versus using one's professional title, such as "Dr." last name) and the instructor's demographics (i.e., race and gender). Thus, it is important to continue building on the current research that examines the influences on and the usefulness of SEIs.

STUDY 1

Instructors often decide how they would like undergraduate students to refer to them; some use their profession title (e.g., "Dr." or "Professor"), whereas others are on a first-name basis with their students. People can treat others differently depending on their name (Watson, Appiah, Thornton 2011); however, little research has examined the honorifics of the name. One study found that graduate students rated faculty members who were addressed by their first name as more approachable and helpful than those who were addressed by their formal title (McDowell and Westman 2005). Another study found that female professors (but not male professors) who use the title "Dr." are perceived to be less accessible (Takiff, Sanchez, and Stewart 2001). That study found evidence for this effect in two different ways. First, when students reported how they addressed their instructors, they more often referred to their male professors with their professional title. In their follow-up study, Takiff, Sanchez, and Stewart (2001) manipulated the instructor's name and gender by having students read a hypothetical transcript of a class. In the transcript, hypothetical students in the class referred to the instructor by either their first name (i.e., "Richard" for male and "Sharon" for female) or as "Professor Parks." Whereas female professors who went by their professional title were perceived as equal in status to male professors who went by their professional title, the former were rated lower in accessibility. These results are concerning for female professors who may feel the need to choose between being perceived as having high status (i.e., using their formal title) or being accessible (i.e., using their first name), both of which are important for their career.

Although the Takiff, Sanchez, and Stewart (2001) study provided interesting and concerning data, there is a need to examine this effect on SEIs from actual courses. Based on the concept of teacher immediacy and the likelihood that students may feel closer to instructors with whom they are on a first-name basis—according to Takiff, Sanchez, and Stewart (2001), they may be seen as more accessible—I predicted that student evaluations for a female instructor would be higher when the instructor introduces herself using her first name:

Hypothesis H1: An instructor will receive higher SEI scores for courses in which she expresses a preference for using her first name rather than her professional title.

METHODS

The following methods were used to examine the relationship between an instructor's name preference and SEIs.

Participants

SEI data are deidentified, aggregated, and publicly available online; therefore, a representative from the Institutional Review Board (IRB) confirmed that a human-subjects review was not necessary. SEI data were examined from 16 sections of a course in persuasive communication (Foster 2022).

Measures

Students were asked to evaluate three dimensions of teaching on a five-point scale: (1) instructor's preparedness, organization of material, and clarity of presentation (i.e., well organized, instructor well prepared, communicated subject matter clearly); (2) rapport and instructor commitment (i.e., instructor interested in teaching, instructor interested in helping students, created learning atmosphere); and (3) students' sense of their own learning (i.e., intellectually stimulating, encouraged independent thinking, learned greatly from instructor), as well as an overall rating. However, the nine items for the different dimensions of teaching loaded onto one factor (Cronbach $\alpha = 0.97$; McDonald $\omega = 0.90$), which was highly correlated with the overall rating ($rs = 0.82$). Therefore, the overall ratings were analyzed.

Procedure

Between 2015 and 2020, the same course was taught 16 times by the same instructor. For 10 classes, the instructor expressed a preference for going by her first name and for six classes (table 1), she expressed a preference for going by "Dr." last name. Sometimes there were two sections of the course in the same semester, one of which was the first-name class and the other was the "Dr."-last-name class. Because these courses were taught back to back, the order of which class (i.e., earlier or later) was first name or "Dr." last name was counterbalanced. All classes were taught in person, and the instructor's name preference was on the syllabus, mentioned on the first day of class, and reinforced with a name tent visible throughout the lecture that listed either her first name or "Dr." last name. Other than the difference in name preference, the classes were designed to be equivalent.

RESULTS AND DISCUSSION

The mean SEI response rate across the classes was 88%. This high response rate likely was due to all course sections being offered extra credit if the class reached an 80% or higher completion rate. As expected, SEIs were higher in classes in which the instructor expressed a preference to be called by her first name ($M = 4.76$, $SD = 0.06$) than in classes in which she expressed a preference to be called "Dr." followed by her last name [($M = 4.37$, $SD = 0.18$), $F(1, 14) = 38.34$, $p < 0.0001$, $d = 3.54$]. Similar results were obtained when the courses were weighted by the class size [$Ms = 4.76$ and 4.37, $F(1, 14) = 37.39$, $p < 0.0001$] and number of responses [$Ms = 4.76$ and 4.37, $F(1, 14) = 35.89$, $p < 0.0001$]. Similar results also were obtained when the response rate was used as a covariate in the model [adjusted $Ms = 4.76$ and 4.37, $F(1, 13) = 27.83$, $p < 0.0001$].

The findings are consistent with previous research on teacher–student rapport (Richmond et al. 2015) and the concept of teacher

Overall Evaluation for Classes with Instructor Using First or Last Name

Year	Semester	Students	Responses	Response Rate	Name	SEI
2015	Autumn	27	22	81%	First	4.90
2016	Autumn	23	19	83%	First	4.80
2017	Autumn	22	22	100%	First	4.70
2017	Autumn	25	24	96%	Dr. Last	4.50
2018	Spring	25	23	92%	First	4.70
2018	Spring	23	18	78%	First	4.80
2018	Autumn	26	24	92%	First	4.71
2018	Autumn	26	23	88%	Dr. Last	4.13
2019	Spring	23	20	87%	First	4.80
2019	Spring	27	22	81%	First	4.73
2019	Spring	27	25	93%	Dr. Last	4.52
2019	Autumn	28	27	96%	Dr. Last	4.52
2019	Autumn	28	26	93%	Dr. Last	4.15
2019	Autumn	25	22	88	Dr. Last	4.41
2020	Spring	26	23	88	First	4.74
2020	Spring	25	18	72	First	4.72

Note: SEI=student evaluation of instruction.

immediacy (Moore et al. 1996), suggesting that students feel closer to instructors who prefer to be called by their first name. From this study, however, it is important that there was a significant difference in SEI scores, not perceived closeness.

One limitation is that the instructor for all courses was a white female. It is important to determine whether these findings replicated for other instructors, especially given the complex relationship between an instructor's gender and SEIs (Rosen 2018; Wong and Bouchard 2021). Previous research suggests that female professors (but not male professors) who use the title "Dr." are perceived to be less accessible (Takiff, Sanchez, and Stewart 2001). Additionally, minority instructors have been found to receive lower evaluations than white instructors (Carle 2009).

STUDY 2

The results of Study 1 provide data from an actual course that are consistent with previous research (Takiff, Sanchez, and Stewart 2001) that used a hypothetical course. However, it is unclear whether these results would be generalizable to other instructors. Additionally, other demographic characteristics of the instructor (e.g., gender and race) could interact with name preference to influence SEIs. Thus, an experimental study explored these additional demographics, which acknowledged that experimental manipulations could provide insight into real-world phenomena but do not replicate them.[1]

Based on the results of Study 1, I predicted more favorable teacher ratings for instructors who expressed a preference for using their first name (i.e., "Brian" for males and "Rachel" for females) than for instructors who expressed a preference for using "Dr. Moore." These names were chosen because they were in the top 16 most common names from 1980,[2] making them plausible names for an instructor.

Hypothesis H1: Instructors who introduce themselves using their first name will have higher evaluations than instructors who introduce themselves using their professional title.

Additionally, gender may influence SEIs and may interact with name preference (Takiff, Sanchez, and Stewart 2001). Thus, I

As expected, teaching evaluations were higher in classes in which the instructor expressed a preference to be called by her first name (M = 4.76, SD = 0.06) than in classes in which she expressed a preference to be called "Dr." followed by her last name [(M = 4.37, SD = 0.18), $F(1, 14) = 38.34$, $p < 0.0001$, d = 3.54].

predicted less-favorable SEIs for female instructors (but not for male instructors) who expressed a preference for using "Dr. Moore."

Hypothesis H2: Female (but not male) instructors who introduce themselves using their first name will have higher evaluations than instructors who introduce themselves using their professional title.

Additionally, the instructor's race may influence SEIs (Carle 2009) and may interact with name preference or gender. For example, one study found that female minority instructors were rated lowest on student evaluations (Chavez and Mitchell 2020). Thus, I predicted that Black female instructors who preferred to use "Dr. Moore" would receive the lowest teacher ratings and white female instructors who preferred to use their first name would receive the highest teacher ratings.

Hypothesis H3: The lowest evaluations will be for instructors who are Black and female and who introduce themselves using their professional title.

I also sought a better understanding of why these demographics and preferences may influence SEIs. Using a mediation model, I predicted that a name preference for "Dr. Moore" would decrease teacher ratings through a decrease in perceived teacher immediacy—in this case, defined by teachers who are perceived as being more "friendly," "engaging," "fun," "nurturing," and "caring"—especially for female instructors. Previous research has shown that students expect female instructors to be more caring (Andersen and Miller 1997; Langbein 1994), and female instructors might be perceived as less caring if they request students to use the title "Dr." when referring to them.

Hypothesis H4: Instructors who introduce themselves using their first name will have higher evaluations because of increased perceptions of being friendly and closer to students and also depending on the instructor's race and gender.

Furthermore, I expected that this mediation might depend on individual student characteristics (e.g., political ideology). However, based on previous research (Mosso et al. 2013; Ratliff et al. 2019), I expected this result to be limited to students who identify as conservative and Republican.

Hypothesis H5: Conservative students will rate Black and female instructors lower on teaching evaluations when the instructors introduce themselves by their professional title rather than by their first name.

METHODS

The following methods were used to examine the relationship between an instructor's demographics and SEIs.

Participants

Study 2 was approved by the university's IRB (Protocol 2021B0044). A power analysis revealed that 512 participants (i.e., 64 in each of the eight conditions) were needed to detect a medium-sized effect ($d = 0.5$) at the 0.05 significance level (two-sided), given that power $= 0.80$ (Cohen 1988). The participants included 648 undergraduate students who received course credit for their voluntary participation. Those who failed the two attention-check items were excluded (N = 150), resulting in a total of 498 valid responses. Thus, the study was slightly under-powered. There were 331 females, 165 males, and two nonbinary participants: 69.5% white, 10.0% Black, 14.5% Asian, and 6.0% Hispanic. Most participants identified as being affiliated with the Democratic Party (45%), followed by Independents (30%) and Republicans (17%); 8% declined to respond. They ranged in age from 18 to 63 ($M = 20.30$, $SD = 3.59$).

Measures

Using exploratory factor analysis, indices were created, as follows:

- Four items measured the likelihood of registering for the course (i.e., the scale with the following stems: "How likely are you to" "register for this class," "recommend this class to a friend," "tell a friend about this class," and "want to learn more about this class") (Cronbach α = 0.88; McDonald ω = 0.88).

- Six items measured evaluations of the course (i.e., whether they thought it was useful, interesting, engaging, fun, intellectually stimulating, and worthwhile) (Cronbach α = 0.87; McDonald ω = 0.87).
- Eight items measured support for traditionally liberal social and political issues including gender equality, racial equality, women's rights, LGBTQ+ rights, Black Lives Matter, universal healthcare, free college, and the environmental movement (Cronbach α = 0.89; McDonald ω = 0.89).
- Three items measured support for traditionally conservative social and political issues including the men's rights movement, All Lives Matter, and Blue Lives Matter (Cronbach α = 0.83; McDonald ω = 0.85).
- Five items measured instructor immediacy (i.e., whether the instructor was friendly, engaging, fun, nurturing, and caring) (Cronbach α = 0.89; McDonald ω = 0.89).
- Nine items measured SEIs as in Studies 1 and 2 (Cronbach α = 0.90; McDonald ω = 0.90).

Procedures

Similar as in the Takiff, Sanchez, and Stewart study (2001), students were told that they would be introduced to a new course to gauge their interest. However, whereas the Takiff Sanchez, and Stewart study had participants read transcripts of the course and manipulated whether a student in the transcript referred to the instructor by their first name or title, this manipulation had the instructor directly share their name preference with students in a welcome video.

The welcome video was between 28 and 31 seconds long and showed the instructor facing the camera from the shoulders up with no background objects (i.e., a plain, neutral background). The script was the same for all instructors except for their name preference, which was either "Dr. Moore," "Brian" for male instructors, or "Rachel" for female instructors. Specifically, they started by stating their full name and then their preference; for example, "My full name is Dr. Brian Moore, but I prefer to go by Brian" or "My full name is Dr. Rachel Moore, but I prefer to go by Dr. Moore." After viewing the welcome video, students read the course syllabus (which differed only in the name of the instructor) and answered survey questions.

To increase generalizability (Wells and Windschitl 1999), I used two examples of each instructor's demographic—that is, there were two white males, two Black males, two white females, and two Black females who created welcome videos, introducing themselves with a preference for either their first name or their professional title. Thus, the study design was a two (instructor name preference: first or last), by two (instructor race: white or Black), by two (instructor gender: female or male) design.

There was no significant SEI difference between the two examples of white females ($p = 0.70$), the two examples of Black females ($p = 0.90$), the two examples of Black males ($p = 0.69$), or the two examples of white males ($p = 0.11$). Therefore, the data from the two examples of each demographic category were combined.

RESULTS AND DISCUSSION

As expected in Hypothesis H1, SEIs were higher for instructors who preferred going by their first name ($M = 4.25$, $SD = 0.57$) than for instructors who preferred going by "Dr. Moore" [($M = 4.11$, $SD = 0.64$), $F(1, 493) = 7.11$, $p = 0.008$, $d = 0.25$]. However,

the gender-by-name-preference interaction expected in Hypothesis H2 was nonsignificant ($p = 0.32$). Thus, the name-preference effect held for both male and female instructors.

Regarding Hypothesis H3, there was a significant three-way interaction among instructor gender, race, and name preference [$F(1, 493) = 3.79$, $p = 0.05$]. I probed the interaction by examining

Black men who introduced themselves as "Brian" had higher SEIs (M = 4.30, SD = 0.81) than Black men who introduced themselves as "Dr. Moore" [(M = 3.95, SD = 0.87), t (492) = 2.67, p = 0.03, d = 0.42].

each group separately. Black men who introduced themselves as "Brian" had higher SEIs ($M = 4.30$, $SD = 0.81$) than Black men who introduced themselves as "Dr. Moore" [($M = 3.95$, $SD = 0.87$), $t(492) = 2.67$, $p = 0.03$, $d = 0.42$]. Name preference did not influence teaching ratings for Black women ($p = 0.64$), white men ($p = 0.73$), or white women ($p = 0.13$).

For explanatory mechanisms, I predicted that name preference for "Dr. Moore" would decrease teacher ratings due to a decrease in perceived immediacy. This hypothesis was partially supported because the mediation model was significant only when the instructor was a Black male. Using Hayes's (2018) PROCESS model 4, I found that name preference was predictive of perceived instructor immediacy ($\beta = -0.25$, $p = 0.04$) and that perceived immediacy, in turn, was predictive of SEI ratings ($\beta = 0.56$, $p < 0.001$). For the overall model, whereas the direct effect of name preference on SEI ratings was not significant ($p = 0.28$), the indirect effect through perceived teacher immediacy was significant ($\beta = -0.14$, 95% CI = $-0.27, -0.0098$). Because the preference for "Dr. Moore" was coded as "1" and the preference for "Brian" was coded as "0," those who preferred "Dr. Moore" had lower perceived immediacy and therefore lower SEI scores (correlations are shown in table 2).

I used Hayes's (2018) PROCESS model 7 to examine potential boundary effects for Hypothesis H4 using moderated mediation, but none of these analyses were significant. Thus, the evidence for bias against Black male instructors who expressed a preference for the title "Dr." remained regardless of social and political factors.

GENERAL DISCUSSION

The literature on the efficacy of SEIs is complicated because some studies show no correlation between instructor demographics such as gender and race (Park and Dooris 2020). However, this current study, aligned with other previous research (Mitchell and Martin 2018; Murray et al. 2020), provides evidence that nonteaching factors can influence SEIs. When research yields conflicting results, it often is the case that there are boundary conditions to an effect. Examining possible boundary effects and adding to the existing body of literature, courses in which the instructor preferred her first name rated her higher (i.e., never lower than 4.70 on a 5.00 scale) than courses in which the same instructor preferred her last name proceeded by "Dr." (i.e., never higher than 4.52 on a 5.00 scale).

There are, of course, alternative explanations for the discrepancy between teaching skills and SEIs. For example, it is possible that instructors with a more negative attitude toward SEIs also may receive lower SEIs (Carlozzi 2018). In this case, however, the instructor's SEIs from Studies 1 and 2 were influenced by nonteaching factors despite the SEIs being at or above average for similar courses taught by other instructors.

One notable difference between the actual SEIs in Study 1 and the experiment in Study 2 is that the white female instructors' name preference in the experiment did not predict SEI scores (i.e., significant prediction was only for Black male instructors). Although further research is necessary to determine the reason for this discrepancy, one possibility is that the current study was slightly underpowered, resulting in the effects of name preference on SEI scores for white female instructors approaching but not reaching significance ($p = 0.13$). In post hoc analysis, I found that although the immediacy scale was not predicted as a whole (i.e., indexed by averaging) by name preference for white females, the individual item "friendly" was higher for white females who preferred the use of their first name ($M = 4.09$, $SD = 0.723$) rather than their professional title [($M = 3.75$, $SD = 1.03$), $t(119) = 2.14$, $p = 0.017$]. Additionally, white females who went by "Dr. Moore" were considered more "strict" ($M = 2.98$, $SD = 1.04$) than those who went by "Rachel" [($M = 2.67$, $SD = 0.911$), $t(121) = -1.77$, $p = 0.079$)]. This led me to consider that it is possible that one approximately 30-second introduction video did not have as strong an impact on participants in the experiment as a

Table 2
Correlations from Study 2

Variable	Title	Gender	Race	Immediacy	SEI	Register
Title	1					
Instructor Gender	0.052	1				
Instructor Race	0.008	−0.024	1			
Immediacy	−0.090*	−0.169**	0.186**	1		
SEI	−0.119**	−0.088	0.100*	0.601**	1	
Register	−0.033	−0.042	0.113*	0.440**	0.484**	1
Course Evaluation	−0.047	−0.025	0.098*	0.551**	0.530**	0.615**

Notes: *Correlation is significant at the 0.05 level (two-tailed); **correlation is significant at the 0.01 level (two-tailed).

semester-long course with an instructor. This demonstrates the importance of examining real-world data in addition to experimental studies (see also Feldman 1993). It also raises the question: If the name preference in a 30-second video significantly influenced SETs for Black male instructors in a hypothetical course, what might it do in an actual course?

If the name preference in a 30-second video significantly influenced teacher evaluations for Black male instructors in a hypothetical course, what might it do in an actual course?

CONCLUSION

As in previous research, SEIs continue to be influenced by factors other than teaching quality. In particular, factors relating to perceived closeness to the instructor (i.e., rating the instructor as "friendly," "engaging," "fun," "nurturing," and "caring") appear to influence evaluations, especially for Black male instructors. It is concerning to note that although the SEI scores were relatively high for Black male instructors who went by their first name, their SEI scores decreased dramatically when they went by "Dr. Moore." This result was consistent even among self-identifying liberal students who were supportive of social issues such as the Black Lives Matter movement. This highlights the need to learn more about how instructor demographics can have a (perhaps largely unconscious) role in SEIs, which can impact their career in important ways. This research also provides evidence to support the need for additional methods to evaluate teaching abilities that are less susceptible to bias than SEIs (McCarthy, Niederjohn, and Bosack 2011).

ACKNOWLEDGMENTS

I thank Brad J. Bushman for his helpful comments on an earlier draft of this article.

CONFLICTS OF INTEREST

The author declares that there are no ethical issues or conflicts of interest in this research. ∎

NOTES

1. The following five hypotheses were preregistered at https://osf.io/6xrcv/?view_only=942do6cfc19944bd9ea2ef78ao9617f8.
2. See www.ssa.gov/OACT/babynames/decades/names1980s.html.

REFERENCES

Andersen, Kristi, and Elizabeth D. Miller. 1997. "Gender and Student Evaluations of Teaching." *PS: Political Science & Politics* 30 (2): 216–19.

Carle, Adam C. 2009. "Evaluating College Students' Evaluations of a Professor's Teaching Effectiveness Across Time and Instruction Mode (Online Vs. Face-to-Face) Using a Multilevel Growth Modeling Approach." *Computers and Education* 53:429–35.

Carlozzi, Michael. 2018. "Rate My Attitude: Research Agendas and RateMyProfessor Scores." *Assessments & Evaluation in Higher Education* 43 (3): 359–68.

Chavez, Kerry, and Kristina M. W. Mitchell. 2020. "Exploring Bias in Student Evaluations: Gender, Race, and Ethnicity." *PS: Political Science & Politics* 53 (2): 270–74.

Christophel, Diane M. 1990. "The Relationships Among Teacher Immediacy Behaviors, Student Motivation, and Learning." *Communication Education* 39 (4): 323–40. DOI:10.1080/03634529009378813.

Clayson, Dennis E. 2018. "Student Evaluation of Teaching and Matters of Reliability." *Assessment & Evaluation in Higher Education* 43 (4): 666–81.

Cohen, Jacob. 1988. "Statistical Power Analysis for the Behavioral Sciences." Second edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Esarey, Justin, and Natalie Valdes. 2020. "Unbiased, Reliable, and Valid Student Evaluations Can Still Be Unfair." *Assessment & Evaluation in Higher Education* 45 (8): 1106–20.

Feldman, Kenneth A. 1993. "College Students' Views of Male and Female College Teachers: Part II—Evidence from Students' Evaluations of Their Classroom Teachers." *Research in Higher Education* 34 (2): 151–211.

Foster, Melissa M. 2022. "Replication Data for Instructor Name Preference and Student Evaluations of Instruction.'" https://doi.org/10.7910/DVN/GT89B6, Harvard Dataverse.

Frymier, Ann Bainbridge. 2013. "Teacher Immediacy." In *International Guide to Student Achievement*, ed. John Hattie and Eric M. Anderman, 425–27. London: Routledge/Taylor & Francis Group.

Hayes, Andrew F. 2018. "Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach." New York: The Guilford Press.

Henson, Alisha M., and Elaine Scharfe. 2011. "Association Between Adult Attachment Representations and Undergraduate Student Course Evaluations." *Teaching of Psychology* 38 (2): 106–109. https://doi.org/10.1177/0098628311401582.

Hoel, Anniken, and Tove I. Dahl. 2019. "Why Bother? Student Motivation to Participate in Student Evaluations of Teaching." *Assessment & Evaluation in Higher Education* 44 (3): 361–78.

Holland, E. Penelope. 2019. "Making Sense of Module Feedback: Accounting for Individual Behaviours in Student Evaluations of Teaching." *Assessment & Evaluation in Higher Education* 44 (6): 961–72.

Lakeman, Richard, Rosanne Coutts, Marie Hutchinson, Megan Lee, Deb Massey, Dima Nasrawi, and Jann Fielden. 2021. "Appearance, Insults, Allegations, Blame, and Threats: An Analysis of Anonymous Nonconstructive Student Evaluations of Teaching in Australia." *Assessment & Evaluation in Higher Education.* DOI: 10.1080/02602938.2021.2012643.

Langbein, Laura I. 1994. "The Validity of Student Evaluations of Teaching." *PS: Political Science & Politics* 27 (3): 545–53.

McCarthy, Maureen A., Daniel M. Niederjohn, and Theodore N. Bosack. 2011. "Embedded Assessment: A Measure of Student Learning and Teaching Effectiveness." *Teaching of Psychology* 38 (2): 78–82.

McClain, Lauren, Angelika Gulbis, and Donald Hays. 2018. "Honesty on Student Evaluations of Teaching: Effectiveness, Purpose, and Timing Matter!" *Assessment & Evaluation in Higher Education* 43:369–85. DOI:10.1080/02602 938.2017.1350828.

McDowell, Joan E., and Alida S. Westman. 2005. "Exploring the Use of First Name to Address Faculty Members in Graduate Programs." *College Student Journal* 39 (2): 353–56.

Mehrabian, Albert. 1966. "Immediacy: An Indicator of Attitudes in Linguistic Communication." *Journal of Personality* 34 (1): 26–34. DOI:10.1111/j.1467-6494.1966.tb01696.x.

Mehrabian, Albert. 1971. "*Silent Messages.*" Belmont, CA: Wadsworth Publishing Group.

Meyerberg, Jenna M., and Angela M. Legg. 2015. "Assessing Professor–Student Relationships Using Self-Report Scales." In *A Compendium of Scales for Use in the Scholarship of Teaching and Learning*, ed. Rajiv S. Jhangiani, Jordan D. Troisi, Bethany Fleck, Angela M. Legg, and Heather D. Hussey, 149–60. Washington, DC: Society for the Teaching of Psychology.

Mitchell, Kristina M. W., and Jonathan Martin. 2018. "Gender Bias in Student Evaluations." *PS: Political Science & Politics* 51 (3): 648–52.

Moore, Alexis, John T. Masterson, Diane M. Christophel, and Kathleen A. Shea. 1996. "College Teacher Immediacy and Student Ratings of Instruction." *Communication Education* 45 (1): 29–39. DOI:10.1080/03634529609379030.

Mosso, Cristina, Giovanni Briante, Antonio Aiello, and Silvia Russo. 2013. "The Role of Legitimizing Ideologies as Predictors of Ambivalent Sexism in Young People: Evidence from Italy and the USA." *Social Justice Research* 26 (1): 1–17. https://doi.org/10.1007/s11211-012-0172-9.

Murray, Dakota, Clara Boothby, Huimeng Zhao, Vanessa Minik, Nicolas Berube, Vincent Larivière, and Cassidy R. Sugimoto. 2020. "Exploring Personal and Professional Factors Associated with Student Evaluations of Tenure-Track Faculty." *PLoS ONE* 15 (6): e0233515. https://doi.org/10.1371/journal.pone.0233515.

Park, Eunkyoung, and John Dooris. 2020. "Predicting Student Evaluations of Instruction Using Decision-Tree Analysis." *Assessment & Evaluation in Higher Education* 45 (5): 776–93.

Ratliff, Kate A., Liz Redford, John Conway, and Colin Tucker Smith. 2019. "Engendering Support: Hostile Sexism Predicts Voting for Donald Trump over Hillary Clinton in the 2016 US Presidential Election." *Group Processes & Intergroup Relations* 22 (4): 578–93. https://doi.org/10.1177/1368430217741203.

Remmers, Hermann H., and George C. Brandenburg. 1927. "Experimental Data on the Purdue Rating Scale for Instruction." *Educational Administration and Supervision* 13:519–27.

Richmond, Aaron S., Majken B. Berglund, Vadim B. Epelbaum, and Eric M. Klein. 2015. "a + (b(1)) Professor–Student Rapport + (b2) Humor + (b(3)) Student Engagement = ((Y)Over-Cap) Student Ratings of Instructors." *Teaching of Psychology* 42 (2): 119–25.

Rosen, Andrew S. 2018. "Correlations, Trends, and Potential Biases Among Publicly Accessible Web-Based Student Evaluations of Teaching: A Large-Scale Study of RateMyProfessor.com Data." *Assessment & Evaluation in Higher Education* 43 (1): 31–44.

Seldin, Peter. 1993. "The Use and Abuse of Student Ratings of Professors." *Chronicle of Higher Education* 39 (46): A40.

Stewart, Barbara, Peter Speldewinde, and Benjamin Ford. 2018. "Influence of Improved Teaching Practices on Student Satisfaction Ratings for Two Undergraduate Units at an Australian University." *Assessment & Evaluation in Higher Education* 43 (4): 598–611.

Stroebe, Wolfgang. 2020. "Student Evaluations of Teaching Encourages Poor Teaching and Contributes to Grade Inflation: A Theoretical and Empirical Analysis." *Basic and Applied Social Psychology* 42 (4): 276–94.

Takiff, Hilary A., Diana T. Sanchez, and Traci L. Stewart. 2001. "What's in a Name? The Status Implications of Students' Terms of Address for Male and Female Professors." *Psychology of Women Quarterly* 25 (2): 134–44. DOI:10.1111/1471-6402.00015.

Valencia, Edgar. 2019. "Acquiescence, Instructor's Gender Bias, and Validity of Student Evaluations of Teaching." *Assessment & Evaluation in Higher Education* 45 (4): 483–95.

Watson, Stevie, Osei Appiah, and Corliss G. Thornton. 2011. "The Effect of Name on Pre-Interview Impressions and Occupational Stereotypes: The Case of Black Sales Job Applicants." *Journal of Applied Social Psychology* 41:2405–20. https://doi.org/10.1111/j.1559-1816.2011.00822.x.

Wells, Gary L., and Paul D. Windschitl. 1999. "Stimulus Sampling and Social Psychological Experimentation." *Personality and Social Psychology Bulletin* 25 (9): 1115–25. DOI:10.1177/01461672992512005.

Wong, Jennifer S., and Jessica Bouchard. 2021. "Are Students Gender-Neutral in Their Assessment of Online Teaching Staff?" *Assessment & Evaluation in Higher Education* 46 (5): 719–39.

Wright, Stephen L., and Michael A. Jenkins-Guarnieri. 2012. "Student Evaluations of Teaching: Combining the Meta-Analyses and Demonstrating Further Evidence for Effective Use." *Assessment & Evaluation in Higher Education* 37 (6): 683–99. DOI:10.1080/02602938.2011.563279.

Youmans, Robert J., and Benjamin D. Jee. 2007. "Fudging the Numbers: Distributing Chocolate Influences Student Evaluations of an Undergraduate Course." *Teaching of Psychology* 34 (4): 245–47.

Yu, Kwok W., Lisa Mincieli, and Nina Zipser. 2021. "How Student Evaluations of Teaching Affect Course Enrollment." *Assessment & Evaluation in Higher Education* 46 (5): 779–92. DOI:10.1080/02602938.2020.1808593.