

**Bias against women academics in student evaluations of teaching:  
Tarring and feathering in academia**

Author

Ronksley-Pavia, Michelle

Published

2023

Book Title

Academic Women: Voicing Narratives of Gendered Experiences

Version

Version of Record (VoR)

DOI

[10.5040/9781350274303.0017](https://doi.org/10.5040/9781350274303.0017)

Rights statement

© 2023 Bloomsbury Publishing. The attached file is reproduced here in accordance with the copyright policy of the publisher. Please refer to the publisher's website for further information.

Downloaded from

<http://hdl.handle.net/10072/419660>

Griffith Research Online

<https://research-repository.griffith.edu.au>

## Chapter 6

### BIAS AGAINST WOMEN ACADEMICS IN STUDENT EVALUATIONS OF TEACHING

### TARRING AND FEATHERING IN ACADEMIA

Michelle Ronksley-Pavia

#### *Prelude*

In this chapter I juxtapose and situate the narrated *realities* of gendered experiences of women teaching in academia against relevant literature relating to the all too conspicuous student experience of course (SEC) and student experiences of teaching (SET) surveys and evaluations; collectively I term these student evaluations. The narrated realities (narratives), that I position alongside the exploration of associated literature are a fusion of women's voices from my own lived experiences of working and being in academia over the last ten years; voices of women heard by me, experienced by me and subsumed by me – a personalized narrative of others' happenings fused with those of my own. The fused protagonists in this composition are Shes and Hers; together, these capitalized pronouns are used as proper nouns to assist in providing voice to the many and varied experiences with which female academics contend, alongside literature pertaining to the gender bias intrinsic to student teaching evaluations. In this way, I aim to enable the narratives to convey the multiplicity of women's experiences with student evaluations, while contrasting these experiences with the findings of empirical research.

The scholarship that informs this approach is interconnected and complex, whereby as a method the writing in this chapter is both process and product. It is, therefore, difficult to tease out the overlapping collectivity of approaches and ingredients that have informed this chapter; they are fragments of autoethnography (e.g. Adams 2015), slices of collective biography (e.g. Gonick and Gannon 2013), segments of fictionalizations of real conversations (e.g. Sallis 1999), part narrative-informed reflections (e.g. Polkinghorne 1996), grazes of collective narratives (e.g. Ovchinnikov 2018), and fractions of collective stories that were never formally documented. Together, in this chapter I have chosen to write differently (Gilmore et al. 2019), not only to challenge canonical ways of writing and presenting scholarship of thoughts and ideas, but moreover in a bid to move forward in

collectivizing thoughts, conversations, stories told and experienced. In doing so I hope together we may better understand the occurrences surrounding academic women and the impacts student surveys can have on us. I thus speak with one voice, informed by many.

In essence, this chapter consists of a series of two main parallel narratives: the empirical literature and the lived-experience narratives (in italicized text). The empirical literature can at times seem disconnected from the realities of how student evaluations impact on female academics. In an attempt to bridge this gap, I seek to write differently (Gilmore et al. 2019) by juxtaposing the italicized narratives alongside an encapsulated typological literature review. It is the reader's choice as to how to best connect the parallel narratives of embodied experiences and empirical literature for their own understanding, like a work of art there is no one *correct* interpretation, but many. I deliberately do not seek to deduce, as it is not my aim to analyse any of these narratives but rather to present the *truths* of each and allow the reader to interpret; bringing their own unique background and experiences to this meaning-making process. The parallel narratives can be read together as they flow through the chapter or read separately; as such, only reading the empirical literature sections, disconnected from the italicized lived-experience narratives. Next, the reader may choose to read the italicized narratives separately. Essentially, I provide three ways to read this chapter: (1) as it is set out and presented, (2) reading only the text in standard font and (3) reading only the text in italicized font.

The purpose of this chapter is to recognize and call out the ingrained bias against women that continues to be upheld in university teacher surveys, ratings and evaluations across the globe (Boring et al. 2016). I acknowledge that student evaluation data are profoundly defective and biased against women (Boring 2020; Heffernan 2021; Hornstein 2017; Langbein 2008) – cisgender, transgender, non-binary, gender-neutral, agender, gender-fluid and gender-queer individuals are likely impacted by the gender bias evident in these types of surveys. However, in this chapter I aim to focus predominantly on the fundamental bias against female academics.

### *Prologue*

As far back as 1971 (McKeachie and Lin), it was readily recognized that higher education evaluations of teacher effectiveness by students were fundamentally affected by gender bias against female academics. Indubitably, research spanning more than fifty years continues to identify that teacher ratings are biased against female academics (see for example, Kishler Bennett 1982; Heffernan 2021). Student evaluations are also inherently prejudiced by classism, racism, sexism, homophobia, ageism and ableism and likewise subject to influences of other biases and prejudice. Yet, universities across the globe continue to rely on these flawed data to assess and rate so-called teaching quality of their academics, and indeed continue to require evaluation data for tenured, non-tenured and

precariously employed academics. Moreover, these surveys play increasing roles in career advancement (or not), in terms of evaluating who is promoted, hired and dismissed (Heffernan 2021).

The reference to tarring and feathering in the title of this chapter by extension refers to severely and devastatingly criticizing an individual, to harshly shame or punish a person (Farlex Inc. 2021). Throughout history, tarring and feathering has long been associated with mob *punishment* against individuals to intimidate and humiliate them as perceived *enemies* (Bell 2013); the *Othering* of individuals. I contend that student evaluations of teaching are the modern-day equivalent simulation of tarring and feathering, where women academics in particular are the *victims* of this castigation because they are perceived by some students to have slighted them, or because students respond to some other conscious or unconscious bias through their responses to teaching evaluations. As Stark and Freishtat (2014) contend, ‘anger motivates people to action more than satisfaction does’ (1). Retribution for some students comes in the form of their survey responses that often humiliate and unfairly criticize female academics and their teaching practice because, for example, they were perceived to have not given *good* grades. Of course, I acknowledge that this also happens to some male academics and those of different genders, but the focus of this chapter is on the systemic bias against women in teaching evaluations.

It is empirically (and anecdotally) well recognized that student evaluations of teaching do not measure teaching *effectiveness* (Borch, Sandvoll and Torsen 2020; Boring et al. 2016; Spooren, Brockx and Mortelmans 2013) and that they are fundamentally flawed (Fan et al. 2019; Heffernan 2021).

*She was trying to hold everything together . . . the session had gone well until the discussion turned to the assessment, students had not received their grades well . . . it was like an invisible switch had been flicked and the mob-mentality had materialized! She shifted uncomfortably, standing at the front of the lecture hall, a sea of angry faces, doing Her best to respond to students’ concerns . . . finally the session came to an end. . . .*

*But She couldn’t yet breathe a sigh of relief; a group of angry students remained, about a dozen, some lingering in the background, not quite sure if they wanted to be part of the pack, lingering near the exit, so they could make a quick getaway just in case the situation deteriorated. No such opportunity for Her!*

*The students milled around, surrounding the dais . . . the demanding envoys soon emerged . . . ‘Why did our assessment feedback say this. . .’, ‘Why did I get this mark? I worked so hard! I put in so much effort!’, ‘The marking was too harsh’ . . . She offered to meet with students individually to address their concerns, but She wasn’t going to respond to such a barrage . . . as the group eventually started to dissipate, She caught a couple of under-the-breath comments . . . ‘We’ll get you in the student evaluations!’, ‘We know how much the university relies on those to evaluate staff!’ . . . How could She address that?*

*She retreated to the relative safety of Her office, struggling to hold back the almost inevitable flow of tears. . . . She was not going to allow Herself to become upset . . . She’d completed Her whole doctorate without shedding a tear.*

Bias against female academics is particularly prevalent when students have a tendency to evaluate the person rather than their actual teaching (Fan et al. 2019). Furthermore, time and again existent literature asserts that the widespread use of these ineffective *measures* by universities actually encourages mediocre teaching and inflation of student grades (Isely and Singh 2005; Langbein 2008; Stroebe 2020). The fact that student evaluations are used in ways that encourage grade inflation means that students hold the balance of power over the teaching of academics; thus, students can shape the teaching behaviour of academics in detrimental ways. As Stroebe (2020) points out, this type of systemic enforced influence on these evaluations, where students need good grades and academics need good student evaluations, can be used by students to ‘reward lenient-grading instructors who require little work and to punish strict-grading instructors’ (284). Coupled with the inherent gender bias against women in student evaluations, these types of appraisals have very little validity (if any at all), but their prejudicial outcomes continue to be used in authoritative and iniquitous ways by institutions.

Interestingly, the marketization and commercialization of higher education have seen students become akin to customers, and as Harris and González (2012) suggest, ‘students have adopted that attitude [of being customers] for themselves, coming to higher education to buy the commodity of credential, rather than to learn’ (6). The authors suggest that this distinction, this systemic shift from students as learners to students as customers, will intensify gender bias against women through such avenues as student evaluations. This can be especially marginalizing for women who also challenge embedded stereotypes of what it means to be a female academic with intersecting identities of disability, class, race, gender and so forth; ‘if academic women become service workers who must please, rather than educate, their students, their career advancement will likely be determined to a greater extent than before by their ratings on “customer service” evaluations’ (Harris and González 2012: 6). In this way, further reducing student ratings of teaching to mere ratings of ‘customer service’ makes such evaluations even more meaningless.

The humiliating and simulated *tarring and feathering* of female academics by cohorts of students who make mass-informed judgements about teaching is deplorable. These judgements can be coordinated in person and via social media, where students mass together to respond to perceived injustices; in some respects, this can be akin to mob behaviour, or deindividuation. The social identity model of deindividuation effects proposed by Reicher, Spears and Postmes (1995) is highly relevant here, where social identity performance stemming from students’ perceptions of belonging to their social group can be quite persuading when it comes to instigating others into adopting negative views of teachers and their teaching, subsequently influencing negative evaluations of teaching, in mob-like behaviour with apparent lack of individual responsibility for their own actions. This mob mentality ensues with the adoption of emotional responses impacted by cohort mentality, instead of making rational evaluations based on individual experiences of teaching and university courses. It can be argued that the mere anonymity of teaching evaluations further perpetuates this power imbalance,

which cohorts of students can hold over academics, with academics having no voice or recourse against prejudicial and, at times, personally assaultive remarks evident in these evaluations.

*Sitting uncomfortably perched on the edge of Her office chair, as She knew She shouldn't as it wasn't a position conducive to good posture. . . . She re-read the anonymous student survey comment, just to double-check She hadn't initially misread or potentially misunderstood. . . . 'My professor was totally wrong when She said that I had incorrectly used [the phrase] in my assignment. She needs more experience in [the world of work], She's dumb! Some feedback provided on my assessment was not necessary and came across rude. There also seems to be a misunderstanding by the professor of how this [phrase] works in the real world.' What?!! . . .*

*She had some idea of who this student was because She recalled this particular assignment and how one student had gone way off track. Maybe it would be better in future not to provide so much feedback to students on their assignments? Cross-referencing this qualitative comment with the quantitative ones on the survey, She could see this same student had been the only one to respond that the assessment expectations were not clear! This had brought down the overall Mean score, but then only eight of the sixty-two students had actually completed the survey! What to do with this feedback for next semester? . . .*

### *Interlude*

A recent report by the United Nations Development Programme (2020) stated that almost 90 per cent of people across the world are biased against women. Within academia globally, 43 per cent of all academics identify as women (The World Bank 2020); however, for a variety of complex reasons, in some faculties this percentage is much lower (e.g. sciences). Nevertheless, what many higher education institutions have in common is that they all use some form of teaching evaluations to assess the teaching of faculty members. Furthermore, there is a large body of research spanning almost fifty years that recognizes the gender bias against female academics inherent in teaching evaluations (Mengel, Sauermann and Zolitz 2019). This body of evidence provides substantiation of the impact of bias against women in these ratings of female academics, including significant impacts on mental health and career progression (Heffernan 2021); frequently, academics are left on their own to cope with the results of student evaluations and derogatory qualitative remarks. Some reasons put forward in the literature for gender bias in evaluations relate to aspects such as conscious and unconscious bias from students, students questioning the teaching competence of female academics, and their apparent 'lack of confidence' (Mengel, Sauermann and Zolitz 2019: 560), or appearance of being 'more shy or nervous' (560), traits purportedly associated by some students with female academics.

*Her heart pounded in her chest as She downloaded the student teaching evaluations . . . trembling, Her hand moved the mouse . . . it was the same each semester, sweating palms, pounding chest – dreading the outcome of these anonymous*

student evaluations. Worst of all, were the comments about Her appearance – one student had once commented that She should not wear a blouse that made her bra visible. Another . . . ‘Some make-up wouldn’t go astray’ . . . Yet another . . . ‘What’s with Her accent? She should speak proper English or go home!’ . . . She needed to have some semblance of resilience to keep on coming back semester after semester! The feedback certainly didn’t represent the whole cohort, but it was always those particularly cruel ones which remained with Her semester after semester. What was **WRONG** with Her? No one else talked about receiving these kinds of comments in their evaluations; She must be the only one this was happening to . . . She tried to contain her shaking body, leaning against the desk for some kind of sympathetic anthropomorphic support, in an attempt to halt Her embodied reaction . . .

A study in the Netherlands of about 20,000 student evaluations over a five-year period (2009–13) found that female teachers received scores thirty-seven percentage points lower than male teachers (Mengel, Sauermann and Zolitz 2019). Boring (2017) analysed over 22,000 teaching evaluations conducted during a five-year period at a French university and found that male students favoured male academics and perceived these male academics as having stronger leadership skills and being more knowledgeable than female academics. Moreover, an experimental American study explored an online course where academics assumed identities of avatars with different genders. In this study, the academic with the male avatar was rated considerably higher than the academic with the female avatar, irrespective of the actual gender of the academic (MacNell, Driscoll and Hunt 2015). Taken together, these studies provide further evidence about the inherent bias against women academics clearly evident in teacher evaluations.

*It was time to review the student evaluation results for Her teaching. Opening the spreadsheet, the quantitative results suggested a particularly low score this semester, 2.8 out of 5. Reading on . . . scrolling down . . . She should ‘wear a dress more often’, was one comment that made Her sit up abruptly in Her chair. Should She read on? How did **that** comment remotely relate to Her teaching competence? She wondered, trying not to get too demoralized.*

*She read on . . . ‘In my previous time at University, I was always told by my professors that a large percentage of failing or low grades among students points to a failure on the professor’s part, and therefore the professor should bear responsibility for my low grade.’ She felt this one cut deep . . . the assessments were set by the Certifying Agency; She had no control over these. Maybe the next comment would be more positive . . . more rewarding for Her investment in the semesters teaching . . . Hypnotized by some kind of perverse vacillation between morbid curiosity, visceral loathsomeness at the sight of a vehicular crash scene, and somehow hoping and praying the rest wasn’t as repugnant as it appeared . . . Feeling a lump developing in Her throat, would She ever be able to swallow properly again? She read on . . .*

*Then came the passive aggressive contradictory ones, ‘She has wonderful depth of knowledge, but I do not believe She is a good fit for this course, She lacks the ability to successfully transmit knowledge to Her students’. She felt physically sick! . . . Who could She talk to about how She was feeling? . . . Her supervisor? . . . No, what if she was unaware of the evaluation outcomes? By bringing it to her attention She risked*

*being outed as incompetent! She had to deal with it on her own . . . She internalized it, once again . . .*

Studies have clearly demonstrated that had women academics been male they would have received higher scores in student evaluations; furthermore, that male teachers who were less effective than their female counterparts received higher student evaluations (Boring 2017; Boring et al. 2016; MacNell, Driscoll and Hunt 2015). The extensive literature on bias against women in these types of evaluations can be summed up by Neath's (1996) ironic number one tip to academics wanting to improve their teaching evaluations without actually improving their teaching – 'BE MALE!' Neath (1996) advises female academics to not be too 'demanding' (1365) of their students lest they be more critical on survey questions pertaining to evaluating their availability, their grading and creation of engaging content, metaphorically tarring and feathering their female teachers. Tongue-in-cheek, Neath (1996) advises female faculty to study the 'Bem Sex-role Inventory [to] learn how to be less feminine and more androgynous' (1365) so their ratings will improve!

The expectations for women in academia rarely relate to the actual subject and content of survey questions but instead relate to how women in academia are perceived, and to expectancies that they are supposed 'to be nice, caring and good-looking. Depending on their age, female professors are seen as "girlfriend" or "mother" and not necessarily as professionals' (Wagner, Rieger and Voorvelt 2016: 80). This further reproduces gender stereotyping in relation to aspects of so-called likeability and competence of females. Furthermore, these perceptions and expectations reflect the inherent bias against women across the globe recognized in the United Nations Development Programme report (2020). Yet, these significant shortcomings of student evaluations remain relatively unaddressed by universities. Indeed, typically, academics are presented with some false semblance of *control* over student evaluations, invited to set up the surveys and add questions before they are released to students to complete. However, in reality, most evaluations consist of five mandatory questions that cannot be changed or altered in any way. In particular, evaluation questions ask, 'How satisfied [you the student] is with the teaching of this staff member?' and 'How could this staff member's teaching be improved?'; with the inherent (yet discounted) implication of these types of questions being that students are somehow *qualified* to evaluate these aspects of an academic's work. Borch, Sandvoll and Torsen (2020) quite rightly attempted to explore what it might mean to students to be *satisfied* with an academic's teaching, with their learning activities and so forth, and found that for the most part, students were confused about what they were actually meant to be evaluating – 'the instructor's ability to make the course exciting, the pedagogical approach, the course literature or learning activities?' (90).

*She hated teaching evaluation results time! There was always this one (male) professor who gleefully shared the results of his evaluations on every social media platform he could find . . . 100% across all the questions, and scores of five out of five on each question! Time and again, he's nominated by his students for teaching awards . . . He's popular among students, he makes jokes, his lectures are just like a stand-up*

*comedy show! He's really entertaining! He's well-known for giving good grades . . . It's all hearsay, but he's rumoured to ask his Teaching Assistants to increase his students' grades when marking . . . But She couldn't help but compare Herself to this evaluation benchmark; She'd never received 100% on any one question, let alone across them all! How does One even do that? Maybe She should try stand-up comedy?! Or give good grades all the time irrespective of work quality?! Sell Her soul to the academia 'Devil'?!*

### *Apologue*

We know that academia is a very competitive environment; academics frequently compete against each other within and across institutions for funding, awards, tenure and promotions. Student evaluations play a large role in specific aspects of an academic's competitiveness. Requirements vary from institution to institution, but generally some form of student evaluation is a requirement for promotion applications, usually across a sequence of semesters of teaching. It can be said quite equivocally that student evaluations of teaching have a persuasive impact on career progression (or not) (Mengel, Sauermann and Zolitz 2019). Furthermore, perceived low scores on survey questionnaires require responses from those academics teaching the particular low-scoring courses. This can be in the form of a comprehensive plan required to *improve* the course, as a direct result of its low score on the flawed evaluation surveys, which takes considerable effort on the part of an academic, often unsupported by university peers or management. This has been found to be particularly problematic for early-career female academics who may be required to put their efforts more into their teaching to the detriment of their research, in turn leading to fewer and potentially lower quality research outputs (Mengel, Sauermann and Zolitz 2019). The resultant course improvement process can be isolating, inducing feelings of failure and impacting on self-confidence, feeling the brunt of being ostracized due to being tarred and feathered.

*It was only Her second year of university teaching, Her third semester since completing Her doctorate . . . She was so confident when opening the evaluation results that they would be good; She had invested so much into the course . . . substantial resources – videos unpacking concepts, support resources and study sheets . . . weekends spent away from Her two young children and Her partner . . . missing family outings and events . . . She'd even missed the school nativity play . . . She'd worked through the immense guilt of those missed events . . . the payoff; She told herself, would eventually be tenure. . .*

*She found the evaluations depressing, not at all what She had expected and secretly hoped for . . . The overall score was 2.5 out of 5 . . . But the focus was on that and that alone, no recognition of the many positive comments in students' qualitative feedback . . . no recognition that this was only the second semester this relatively new course had been offered . . . Actually, Her first time teaching this course!*

*Just Her supervisor's focus on **that** question which asked, 'Overall I am satisfied with the teaching of this academic'. Even She knew that 'Satisfaction' did not equate to quality but was so very subjective! Almost 20 per cent of the 350 students had responded in the survey, a good response! But on that one 'quality' question, 6 per cent of students had selected either 'Dissatisfied' or 'Strongly Dissatisfied', coupled with those who were 'Neutral' (which somehow contributed?).*

*Now She was required to create a mandatory Course Improvement Report . . . She discovered the previous (female) academic who had since left the university, had also had to produce one of these Reports . . . But it had never been passed on to Her, about potential ways to 'improve' the course! She painstakingly compiled the requisite Report . . . dutifully submitting the Report juts prior to the due date, in doing so sacrificing valuable research and writing time . . .*

*The next semester, the course was taken from Her and given to a professor . . . a male professor . . . She felt physically sick; She asked 'Why?' only to be met with platitudes . . . She cried, Her tears were tears of exasperation, of despair, of anger and of frustration . . . no chance to implement Her plan, no opportunity to improve and demonstrate how She had implemented student feedback and made changes to the course . . . Was She such a failure as a teacher that they had to give the course to someone else . . . a man? She grieved for the lost time, the lost opportunity, the missed research and writing deadlines, the unmet sacrifices . . .*

According to Mengel, Sauermann and Zolitz (2019) there is a 'sizeable and systematic bias' (537) against female academics in student evaluations. They claim that this inherent bias affects female academics' careers in three main ways: (a) when evaluated, females appear on the surface to be systematically worse at teaching when compared to males; (b) women academics' publication outputs and publication track records are being negatively impacted by a forced focus on teaching-related activities and (c) the teaching evaluation 'gender gap' (Mengel, Sauermann and Zolitz 2019: 537) may impact on women's self-confidence and beliefs about their ability to teach. The combined impact of these three main factors may be why more women than men leave academia after graduate school (Mengel, Sauermann and Zolitz 2019). Commonly, institutions hold expectations that academics must achieve a rating of 3.75 out of 5 on surveys (Heffernan 2021), otherwise they need to show how teaching will be improved by compiling and writing a report or some form of a plan. Yet, 3.75 out of 5 is 75 per cent; even students receive a passing grade on 50 per cent for most of their assessments, so the question that needs posing is: Why are the surveys not set at 50 per cent (i.e. 2.5 out of 5)? If academics required all students to receive 75 per cent for their assessments, the pass rate and retention rate would likely fall, and universities would likely lose fees from students who could not achieve this level. Yet, academics are expected to achieve 75 per cent semester after semester. As Stark and Freishtat (2014) contend, some universities deliberately instigate 'continuous cycles of competition amongst academics by making the acceptable result one that is above the cohort's average' (Heffernan 2021: 3). In other words, irrespective of the results of the cohort, there will be 50 per cent of academics who are below the average (Stark and Freishtat 2014). Acknowledging the fundamental bias against female academics in these

evaluations remains unstated in such competitive environments, and women continue to be impacted by these institutionalized misogynistic practices.

### *Epilogue*

In academia globally, we have reached a stage of critical mass when it comes to the tide and groundswell that unequivocally confirms the prevalence of prejudice against women inherent in student evaluations. This chapter has unambiguously shown that this groundswell is based on a solid foundation of empirical research. Now is the time for critical change; our tipping point – our ‘moment of critical mass, the threshold, the boiling point’ (Gladwell 2000: 12) for institutional change has been reached.

Across universities, and within a broad range of topic areas taught in these institutions, student evaluations have been shown to better measure student bias against women than they actually measure teacher effectiveness (Boring et al. 2016). It is well recognized that student evaluations disadvantage female teachers in academia. The false belief in the quality of student evaluation data is rampant in universities, yet what remains unseen and unacknowledged by universities is the ingrained bias and prejudice ‘that shape the views that form the [so-called objective] data’ (Heffernan 2021: 3). Discriminatory student evaluations should not be used by institutions to make decisions about whom to hire, fire and promote, as this further contributes to institutionalized practices that marginalize female academics. Research has shown that gendered prejudice in student evaluations varies from course to course and from university to university, and as such the responsibility should be on universities to show that relying on these evaluations does not have a detrimental and inequitable bearing on female academics and other under-represented people (Boring et al. 2016). Universities need to introduce policies that address ingrained prejudice against women in student evaluations and ratings of teaching (Hoorens, Dekkers and Deschrijver 2021) or, better yet, get rid of them all together. But then what would they be replaced with? Better the *Devil* you know?

Gendered inequalities intrinsic in student evaluations are rendered invisible by the gender-blind systemic implementation of these surveys for evaluating the teaching and pedagogical practices of female academics. It is past time that universities recognized this ingrained bias in student evaluations. Transformation and initiation of structural change in the implementation and use of student evaluations is vital to support women academics, indeed all academics, in being the best teachers they can be, while being fully and wholeheartedly supported by their institutions in ways that do not turn a blind eye to misogynistic teaching evaluations. These forms of denial and lack of institutional recognition of inherent gender bias in teaching evaluations are suggestive of institutionalized *cultures of denial* (Cohen 2001), where the perpetuation of patriarchal practices persists through higher education practices and processes, despite apparently progressive institutional policies and initiatives purportedly taken to prevent

such practices. The discernible gender bias in student evaluations has been evident for over fifty years. Women in academia contend with more than enough invisible barriers in endeavouring to achieve equity in the workplace. It is long past time to remove this visible, yet institutionally unacknowledged barrier, in the form of student evaluations, to strive to reach more equitable forms of gender parity in academia.

### References

- Adams, T., S. L. Holman Jones and C. Ellis (2015), *Autoethnography: Understanding Qualitative Research*, New York: Oxford University Press.
- Bell, J. L. (2013), '5 Myths of Tarring and Feathering', *Journal of the American Revolution* [Online]. Available online: <https://allthingsliberty.com/2013/12/5-myths-tarring-feathering/> (accessed 1 October 2021).
- Borch, I., R. Sandvoll and R. Torsen (2020), 'Discrepancies in Purposes of Student Course Evaluations: What Does It Mean to be "Satisfied"?', *Educational Assessment, Evaluation and Accountability*, 32: 83–102. Available online: <https://doi.org/10.1007/s11092-020-09315-x> (accessed 1 October 2021).
- Boring, A. (2017), 'Gender Biases in Student Evaluations of Teaching', *Journal of Public Economics*, 145: 27–41. Available online: <https://doi.org/10.1016/j.jpubeco.2016.11.006> (accessed 20 November 2021).
- Boring, A. (2020), 'Surveying Student Evaluations of Teaching: Vital Tool or Flawed Methodology – What do you Think?', *London School of Economics* (blog). Available online: <https://blogs.lse.ac.uk/impactofsocialsciences/2020/04/09/surveying-student-evaluations-of-teaching-vital-tool-or-flawed-methodology-what-do-you-think/> (accessed 1 October 2021).
- Boring, A., K. Ottoboni, P. B. Stark and G. Steinem (2016), 'Student Evaluations of Teaching (Mostly) do not Measure Teaching Effectiveness', *ScienceOpen Research*. Available online: <https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1> (accessed 1 October 2021).
- Cohen, S. (2001), *States of Denial: Knowing About Atrocities and Suffering*, New York: Wiley.
- Fan, Y., L. J. Shepherd, E. Slavich, D. Waters, M. Stone, R. Abel and E. L. Johnston (2019), 'Gender and Cultural Bias in Student Evaluations: Why Representation Matters', *PLoS ONE*, 14 (2). Available online: <https://doi.org/10.1371/journal.pone.0209749> (accessed 1 September 2021).
- Farlex Inc. (2021), 'Tar and Feather', *The Free Dictionary of Idioms*. Available online: <https://idioms.thefreedictionary.com/tarred+and+feathered> (accessed 1 December 2021).
- Gilmore, S., N. Harding, J. Helin and A. Pullen (2019), 'Writing Differently', *Management Learning*, 50 (1): 3–10, doi: 10.1177/1350507618811027
- Gladwell, M. (2000), *The Tipping Point: How Little Things can Make a Big Difference*, New York: Little Brown and Co.
- Gonick, M and S. Gannon (2013), 'Collective Biography: An Introduction', *Girlhood Studies*, 6 (1): 7–12.
- Harris, A. P. and C. G. González (2012), 'Introduction', in G. Gutiérrez y Muhs, Y. Flores Niemann, C. G. González and A. P. Harris (eds), *Presumed Incompetent: The*

- Intersections of Race and Class for Women in Academia*, 1–14, Boulder, CO: University Press of Colorado.
- Heffernan, T. (2021), 'Sexism, Racism, Prejudice, and Bias: A Literature Review and Synthesis of Research Surrounding Student Evaluations of Courses and Teaching', *Assessment & Evaluation in Higher Education*, 46 (6): 1–11. Available online: <https://doi.org/10.1080/02602938.2021.1888075> (accessed 1 December 2021).
- Hoorens, V., G. Dekkers and E. Deschrijver (2021), 'Gender Bias in Student Evaluations of Teaching: Students' Self-affirmation Reduces the Bias by Lowering Evaluations of Male Professors', *Sex Roles*, 84: 34–48. Available online: <https://doi.org/10.1007/s11199-020-01148-8> (accessed 1 December 2021).
- Hornstein, H. A. (2017, January 1), 'Student Evaluations of Teaching are an Inadequate Assessment Tool for Evaluating Faculty Performance', *Cogent Education*, 4 (1): 1–8. Available online: <https://doi.org/10.1080/2331186X.2017.1304016> (accessed 1 October 2021).
- Isely, P. and H. Singh (2005), 'Do Higher Grades Lead to Favorable Student Evaluations?' *Journal of Economic Education*, 36 (1): 29–42. Available online: <https://doi.org/10.3200/JECE> (accessed 1 June 2021).
- Kishler Bennett, S. (1982), 'Student Perceptions of and Expectations for Male and Female Instructors: Evidence Relating to the Question of Gender Bias in Teaching Evaluation', *Journal of Educational Psychology*, 74 (2): 170–9.
- Langbein, L. (2008), 'Management by Results: Student Evaluation of Faculty Teaching and the Mis-measurement of Performance', *Economics of Education Review*, 27: 417–28. Available online: <https://doi.org/10.1016/j.econedurev.2006.12.003> (accessed 1 July 2021).
- MacNell, L., A. Driscoll and A. N. Hunt (2015), 'What's in a Name: Exposing Gender Bias in Student Ratings of Teaching', *Innovative Higher Education*, 40 (4): 291–303. Available online: <https://doi.org/10.1007/s10755-014-9313-4> (accessed 1 October 2021).
- McKeachie, W. J. and Y. Lin (1971), 'Sex Differences in Student Response to College Teachers: Teacher Warmth and Teacher Sex', *American Educational Research Journal*, 8 (2): 221–6.
- Mengel, F., J. Sauermann and U. Zolitz (2019), 'Gender Bias in Teaching Evaluations', *Journal of the European Economic Association*, 17 (2): 535–66. Available online: <https://doi.org/10.1093/JEEA> (accessed 21 June 2021).
- Neath, I. (1996), 'How to Improve Your Teaching Evaluations Without Improving Your Teaching', *Psychological Reports*, 78 (3\_suppl): 1363–72. Available online: <https://doi.org/10.2466/pr0.1996.78.3c.1363> (accessed 11 June 2021).
- Ovchinnikov, F. (2018), 'Collective Narrative Methodologies'. Available online: <https://fofchik.medium.com/collective-narrative-methodology-draft-40b99bd16151> (accessed 1 June 2022).
- Polkinghorne, D. E. (1996), 'Explorations of Narrative Identity', *Psychological Inquiry*, 7 (4): 363–7.
- Reicher, S., R. Spears and T. Postmes (1995), 'A Social Identity Model of Deindividuation Phenomena', in W. Stroebe and M. Hewstone (eds), *European Review of Social Psychology*, 6: 161–98, Chichester: Wiley.
- Sallis, E. (1999), 'Research Fiction', *TEXT*, 3 (2). Available online: <http://www.textjournal.com.au/oct99/sallis.htm> (accessed 1 June 2022).
- Spooren, P., B. Brocx and D. Mortelmans (2013), 'On the Validity of Student Evaluation of Teaching: The State of the Art', *Review of Educational Research*, 83 (4): 598–642. Available online: <https://doi.org/10.3102/0034654313496870> (accessed 21 June 2021).

- Stark, P. and R. Freishtat (2014), 'An Evaluation of Course Evaluations', *ScienceOpen Research*. Available online: <https://doi.org/10.14293/s2199-1006.1.sor-edu.aofrqa.v1> (accessed 25 June 2021).
- Stroebe, W. (2020), 'Student Evaluations of Teaching Encourages Poor Teaching and Contributes to Grade Inflation: A Theoretical and Empirical Analysis', *Basic and Applied School Psychology*, 42 (4): 276–94. Available online: <https://doi.org/10.1080/01973533.2020.1756817> (accessed 1 October 2021).
- United Nations Development Programme (2020), 'Tackling Social Norms: A Game Changer for Gender Inequalities.' Available online: <https://news.un.org/en/story/2020/03/1058731> (accessed 1 October 2021).
- Wagner, N., M. Rieger and K. Voorvelt (2016), 'Gender, Ethnicity and Teaching Evaluations: Evidence from Mixed Teaching Teams', *Economics of Education Review*, 54: 79–94. Available online: <http://dx.doi.org/10.1016/j.econedurev.2016.06.004> (accessed 21 June 2021).
- World Bank, The (2020), 'Tertiary Education, Academic Staff (% Female)', *UNESCO Institute for Statistics*. Available online: <https://data.worldbank.org/indicator/SE.TER.TCHR.FE.ZS> (accessed 6 September 2021).