Understanding the Impact of Biased Student Evaluations: An Intersectional Analysis of Academics’ Experiences in the U.K. Higher Education Context

**Author details**:

Dr Athanasia Daskalopoulou

Senior Lecturer in Marketing

University of Liverpool Management School

Chatham Street, Liverpool L69 7ZH, United Kingdom

Email: a.daskalopoulou@liverpool.ac.uk

**ABSTRACT**

This paper aims to offer an understanding of how (intersectional) bias in student evaluations of teaching affects academics’ mental health and career progression. Despite the widespread acceptance of student evaluations, an emerging stream of research has begun to highlight the biases and prejudices that underpin much of data collection when it comes to student evaluation surveys both in terms of who completes the evaluation and also in terms of who is being evaluated. While most studies provide analyses of students’ comments and illustrate the abuse that is directed at (marginalised) academics, very little research has focused on the impact of this process on academics. This study offers an overview of academics’ experiences with student evaluations by drawing on an intersectional analysis of 17 interviews with academics employed in U.K. Business Schools. The findings illustrate the detrimental impact of student evaluations on academics’ mental health and career progression. This study also shows how institutional pressures to keep students happy impact academics’ wellbeing and teaching approaches.

**KEYWORDS:** student evaluations; mental health; intersectionality; gender; bias.

Understanding the Impact of Biased Student Evaluations: An Intersectional Analysis of Academics’ Experiences in the U.K. Higher Education Context

**INTRODUCTION**

Student evaluations of teaching are used widely in Higher Education as a means to assess teaching effectives and students’ learning experiences. Universities also use them when making decisions in regard to who is rewarded (e.g., in the allocation of teaching awards, hiring and promotion decisions, etc.) or who is penalised (e.g., in the reallocation of modules, when firing teachers, etc.). Globally over 16,000 Higher Education institutions collect student evaluations (Cunningham-Nelson et al. 2019). In the U.K. Higher Education context student evaluations are the primary means of assessing learning and teaching.

Despite widespread acceptance, an emerging stream of research has begun to highlight the biases and prejudices associated with student evaluations (Constantinou and Wijnen‑Meijer 2022; Heffernan 2021; 2023; Peterson et al. 2019; Utoft and Pradhan 2023). Marsh (2007, 319) argues that student evaluations are “primarily a function of the instructor [..] rather than the course that is taught”. For example, prior research shows that there is a gender bias in student evaluations that impacts women academics more negatively than men (MacNell et al. 2015; Mengel et al. 2017). Additionally, Fan et al. (2019) showed that apart from gender, culture negatively impacts student evaluations of women and teachers of non-English-speaking backgrounds.

Yet, despite having an understanding of the impact of these biases and prejudices, universities continue to use student evaluations. For example, Heffernan’s (2023) study exposed in great detail the detrimental effects of abusive comments to academics’ overall wellbeing and also highlighted universities’ unwillingness to protect their staff. This paper follows Heffernan’s (2023) call for action to further examine how academics are impacted both on a personal level when it comes to their mental health and wellbeing and also on a professional level when it comes to career progression opportunities. As such, this study directly builds on the critical literature around student evaluations (Chávez and Mitchell 2020; Fan et al. 2019; Heffernan 2021; 2023; Mengel et al. 2017; Tucker 2014; Utoft and Pradhan 2023) in order to examine how they affect academics, and who is affected. To do so, an intersectional approach is followed (Collins 2015; Collins and Bilge 2016; Crenshaw 1991) in order to understand how gender, race, (dis-)ability, sexuality, age etc. intersect to produce social locations of (dis-)advantage for academics. This study documents 17 academics’ (working in U.K. Business Schools) emotional, lived experiences with student evaluations across different stages in their careers and shows how different systems of power shape how students perceive academics’ performance. Moreover, this study shows the impact of (biased) student evaluations on academics’ career progression, and also how institutional pressures to keep students happy impact academics’ wellbeing and teaching approaches.

**LITERATURE REVIEW**

**Understanding bias**

Understanding the role of bias in student evaluations is important because they often reflect students’ perceptions about teachers rather than teaching per se (Marsh 2007). Research shows that students often have a high degree of prejudice, albeit implicit, against the person teaching the course (Peterson et al. 2019). For example, Mengel et al. (2017) found that women academics receive systematically lower teaching scores due to gender bias. This study also revealed that a statistical relationship exists between gender bias and ageism; young women receive worst evaluations. Other sources of bias include race and language (Chávez and Mitchell 2020; Johansson and Śliwa 2014; Utoft and Pradhan 2023). For example, Fan et al. (2019) found that gender and culture impact negatively student evaluations of women and academics of non-English speaking backgrounds. However, some studies argue that there is a level of ambiguity on the question of whether the quantitative results in evaluations do, or do not, show clear evidence of lower scores for female academics (e.g., Feldman 1992; 1993; Wright and Jenkins-Guarnieri 2012) indicating the need for more research in this area. At the same time, a relatively small minority of academics are particularly favoured; Heffernan (2021) argues that academics who are white, able-bodied, heterosexual, male, and fall within a specific age range not only face minimal consequences but also gain advantages from the evaluations process.

When considering student evaluations, it is also important to consider who completes the evaluations. For example, within U.K. Higher Education it is common to encounter low response rates, which impact the accuracy of the evaluation (Stark and Freishtat 2014). In addition, results can be further skewed when (some) respondents are motivated by anger (Tucker 2014). Evaluations can also be particularly affected by the demographic characteristics of individual students (Basow 1995; Boring 2017; Heffernan 2021; Mengel et al. 2017). For example, Boring (2017) argues that male students tend to favour male professors; they provide significantly higher scores on overall satisfaction as well as on preparation and organisation of classes, leadership, ability to encourage group work, availability and quality of personal contact, ability to relate to current issues and contribution to intellectual development. Interestingly, Boring (2017) also found that female students favour male professors. Gender concordance is found to be a good indicator of positive student evaluations for male academics in other studies as well. For example, Mengel et al. (2017) illustrated that male students rate female academics 21% worse than male academics. In addition to the effect of students’ gender, Fan et al. (2019) showed that local students give lower scores in evaluations to women and academics from non-English speaking backgrounds than international students (Fan et al. 2019).

Students hold academics accountable to gender specific expectations that are more labour intensive for women (Gelber et al. 2022; Sprague and Massoni 2005). Boring (2017) indicates that male gender stereotypes in relation to authoritativeness and knowledgeability lead male and female students to provide more favourable evaluations to male academics. However, female gender stereotypes in relation to the professor’s ability to be warm and nurturing lead students to provide more similar evaluations to both genders. These findings show that students expect women academics to embody a set of traits and to exhibit a set of behaviours that differs greatly from what is expected from men academics and also students are very critical and penalise women academics that do not conform to these gendered expectations (Gelber et al. 2022). According to MacNell et al. (2015, 10) “male instructors are often afforded an automatic credibility in terms of their professionalism, expertise, and effectiveness as instructors”, whereas women academics are facing a “double-bind”. MacNell et al. (2015) argue that students expect women academics to embody interpersonal traits such as being open and approachable, at the same time these traits are not valued as much as being authoritative and knowledgeable (women violate gendered expectations when they embody these two traits). As such, the “double-bind” refers to the unrealistic standards that women academics need to uphold in order to be valued by students.

**The impact of student evaluations**

From a pedagogical perspective, student evaluations can be harmful for teaching quality, as many academics might shy away from making bold changes out of fear for receiving low scores or might settle for easier assignments/exams or easier to digest course content in order to keep students “happy” (Stark and Freishtat 2014). According to Johnson (2003), academics that grade more leniently are more likely to gain approval and positive evaluations because students ‘reward’ academics that ‘reward’ them. The corelation between positive scores and grades represent a bias because scores are affected by a factor (i.e., grades and/or grade expectations) that is not directly linked to or is an indicator of learning and teaching quality (Johnson 2003).

Student evaluations are widely seen as reliable performance indicators and as such they are used in decision making about (re-)appointing, firing, promoting, or giving tenure to academics (MacNell et al. 2015; Uttl and Smibert 2017). Universities also use them as evidence for shifting research focused contracts to teaching focused, as well as for reallocating research budgets (Mengel et al. 2017). This often affects negatively women and other marginalised and underrepresented academics who are already facing discrimination at the workplace and are thus promoted at a slower pace or are being let go (Heffernan 2021). Bias in student evaluations has also been posited as one of the reasons why women are underrepresented in the upper levels of academic hierarchy and university leadership (Fan et al. 2019; Heffernan 2021). As a result, a plethora of studies argue that student evaluations should not be used for personnel decision making due to their biased nature (Stark and Freishtat 2014).

On a personal level, student evaluations can be a source of anxiety and stress (Stark and Freishtat 2014). As such, they can have a very negative impact on academics’ confidence, especially in the case of junior female academics who tend to receive negative comments as shown by Mengel et al. (2017). Student evaluations are particularly stressful because they often include abusive or unprofessional comments (Tucker 2014); “women academics, academics from different gender and sexual identity groups, and disabled academics receive more [abusive] comments than other groups” (Heffernan 2023, 236). Finally, prejudiced feedback towards faculty members from minority backgrounds can undermine the diversity that Higher Education institutions strive to achieve (Santisteban and Egues 2022). Thus, student evaluations are a major area of concern for staff wellbeing and mental health (Heffernan 2021; 2023).

**METHODOLOGY**

I followed an interpretive qualitative methodology (Denzin and Lincoln 2005) and an intersectional approach to data collection and analysis (Collins 2015; Collins and Bilge 2016; Crenshaw 1991). Following an intersectional approach has enabled me to get a better understanding of academics’ lived-experiences, and the organisation of power within the U.K. Higher Education context by examining how academics’ experiences and responses are “shaped not by a single axis of social division, be it race, or gender, or class, but by many axes that work together and influence each other” (Collins and Bilge 2016, 2).

Participant recruitment involved a mixture of purposeful and snowballing techniques (Patton 2002). The research sample consisted of academics that currently work in U.K. Business (or Management) Schools. I chose to focus on Business Schools as a context due to its increasing internationalisation (Johansson and Śliwa 2014) and because it’s a research context that has been the focus of a number of studies looking at academics’ experiences with sexism and racism (Hughes and Donnelly, 2023). Academics were invited via email to self-select to participate. Initial informants introduced me to additional participants. In total, 17 interviews with academics, lasting up to 90 minutes (average interview 60 minutes), were conducted by the author. My informants aged from 34 to 72 years old, informants self-identified as male (8 informants), female (9 informants), straight (14 informants), bisexual (2 informants) and pansexual (1 informant), informants self-identified as white (5 informants), black (2 informants), South Asian (4 informants), Chinese (2 informants), Arab (3 informants), and African (1 informant). My informants were Assistant Professors (11 informants), Associate Professors (4 informants), and Professors (2 informants), with Teaching and Research contracts (12 informants), Teaching and Scholarship contracts (5 informants), as well as academics who hold positions of responsibility such as Programme Directors and Heads of Teaching (9 informants).

The interviews were audio-recorded on zoom with consent and transcribed verbatim by the author. The interviews explored participants’ lived experiences (Denzin and Lincoln 2005) with student evaluations. I used open-ended, non-directive questions and probes (McCracken 1988 (e.g., how did you feel about the evaluations you have received? How was students’ language/tone in the student evaluations you have received?). During the interview we discussed the feedback they have received, how it made them feel, the impact it has had on their performance and also on their wellbeing and mental health (e.g., have you experienced any positive/negative emotions in relation to student evaluations? Can you please give me an example?). Finally, we unpacked how my informants felt student evaluations have impacted their career progression (e.g., have the evaluations you have received impacted your career progression? If yes, why?).

Pseudonyms were used to protect my informants’ anonymity. All informants received a ‘participant information sheet’ explaining the aims of the research and the interview process and a ‘consent form’ to sign prior to participating in the study. Ethical approval was gained from the author’s institution.

Data analysis was ongoing and followed an inductive approach using thematic coding as I moved back and forth between the data and extant literature (Patton 2002). As such, my findings emerged inductively following a constant comparative logic of analysis to derive my theoretical interpretations (Charmaz 2014). More specifically, I followed the steps of Braun and Clarke’s (2022) reflexive thematic analysis. According to Braun and Clarke (2022), data analysis by a single coder is normal practice. During “dataset familiarisation”, the first phase, the author immersed herself into the data by listening and transcribing the audio files, by (re-)reading transcripts and researcher notes, and by noting initial comments about my informants’ experiences. During the “coding” phase, the author followed a more systematic approach in order to highlight interesting quotations and also allocated meaningful descriptive codes to the dataset. The third phase, “actively generating initial (candidate) themes” involved capturing patterned meaning systematically across all interview transcripts. During the fourth phase of “developing and reviewing themes”, the author compared the initial themes produced in the third phase with the wider research context and extant management education literature. The analysis was fine-tuned during the fifth phase of “refining, defining and naming themes”. Finally, the final phase included the “writing up” of my interpretations.

**FINDINGS**

**The impact of (biased) student evaluations on academics’ wellbeing**

My findings illustrate that going through the process of student evaluations is a mentally draining activity for academics (Heffernan 2023). My informants shared that *“it’s definitely not an objective process”* (George, male, black) and as a result it can be quite stressful for a number of reasons. First of all, academics often have to read derogatory and borderline abusive comments about themselves (e.g., appearance and weight) (Tucker 2014). Sara explains:

***“****I got fat. I got that put on two evaluations. This fat woman, you know. Then just negative comments about what I look like [..] the ones that really stick in my mind are the ones about my weight, [..] it kind of really hurts. And I don’t think students realize that maybe they want to hurt you. What does this fat woman think? What does this fat woman know? That kind of thing, is really quite brutal actually, when you sort of read it. You think? God. [..] There is a level of misogyny in universities, it’s intersectional. I’m older, I’m overweight. I don’t really present as the typical female Professor.”* (Sara, female, white)

Sara’s quotation illustrates how hurtful these comments can be for academics’ mental health. Essentially, students corelated her ability to think and teach with her level of attractiveness (Joye and Wilson 2015; Sebastian and Bristow 2008) and more specifically with being an overweight woman. Sara explains how her disability and age further influence how students perceive her: *“You are registered as disabled when you have [medical condition]. So, it it’s kind of, older woman, working class, disabled. It creates a kind of positionality for you”*. As evidenced in Sara’s narratives, the combination of intersectional biases around gender, class, age, disability and appearance influence how students perceive her as a teacher and how they evaluate her performance. Similarly, Joanna explains:

*“There was more focus on my age. I definitely used to be more conscious of questions of age, and whether people thought I was capable, and how to present [myself] in a way that would come across as being capable.”* (Joanna, female, white)

Joanna, who is in her 30s, reveals the constant effort and energy required in order to compensate for her age. In contrast to Sara, being younger is also a point of contention for students even though it is irrelevant to being a good teacher (Joye and Wilson 2015). Some of my female informants try to combat ageism by *“put[ting] on the lipstick, put[ting] on the heels”* (Jess, female, Indian mixed) whereas some of my young male informants will dress more formally in order to elicit positive comments in student evaluations (Sebastian and Bristow 2008):

*“You would see some kind of biases, whereby if I would start dressing with like a shirt and suit then I would always feel that I get positive comments. If I would dress in a more, let’s say academic way, then you would see that these comments no longer appear, right?”* (George, male, black)

George explains why he chooses to dress more formally. From his experience, he has observed that as a black man in his 30s, he receives more positive evaluations when he wears formal attire. In line with Sara’s, Joanna’s and Jess’ experiences, age here is an important bias which along with race, class, gender, disability, etc. informs students’ perceptions of academics’ identity. As such, these intersecting factors gain different meanings in relation to each other (Collins and Bilge 2016) and produce different experiences for these academics.

My analysis also reveals that my informants that were from non-English speaking backgrounds and from middle and lower-class English-speaking backgrounds received negative comments about their accents in student evaluations in line with other empirical studies (Johansson and Śliwa 2014; Santisteban and Egues 2022). Jess discusses her accent:

*“Even the way that I talk. I’m worried about sometimes. I don’t want to sound uneducated. I don’t want to sound like some angry, arrogant, kind of black woman, or whatever. My accent is not associated with being well-educated, or being high class or being other things that an academic institution should necessarily look like. For me, it has been a real challenge, and it’s an ongoing thing, and it’s something that I have internalized a lot, and I fear sometimes that I’m losing parts of my identity because of it. I don’t want to put on like a fake* *accent, or whatever, like some British accent.”* (Jess, female, Indian mixed)

In this quote Jess reveals the severe impact of these comments for academics from marginalised backgrounds. Indeed, when working within the predominantly white, middle-class, and privileged field of Higher Education (Heffernan 2023), marginalised groups of academics, like Jess in this example, may be more likely to be called to tone-down or even change altogether certain aspects of their identities such as their accents in order to fit in and to receive more positive evaluations.

My findings reveal that student evaluations cause academics to experience a lot of stress and anxiety at the end of each teaching period, even before they can access the data (Heffernan 2023); *“around June every year, I’m just expecting that email with the evaluations. So of course, it’s stressful”* (Olivia, female, white). Olivia explains that every year she is dreading to receive the email with the results of the evaluation despite having worked in Higher Education for ten years. Youssif explains further:

*“When I first started, I felt that everyone needed to be satisfied. I needed to be perfect. No mistakes should be made, and if someone says something negative, then it means that I’m a failure. I’m not good enough. I shouldn’t be teaching.”* (Youssif, male, Arab)

Youssif’s narrative illustrates the additional pressure for Early Career Researchers (ECRs). My informants expressed that at the beginning of their careers they experienced every negative comment as confirmation that they do not belong. As evidenced in Youssif’s quotation, this process intensified he’s imposter syndrome (Addison et al. 2022). Similarly, Mark shares his recollection:

*“It takes one comment to kind of question yourself and your own career. You think like, why do I bother?”* (Mark, male, Arab)

Mark’s quote shows his frustration (Stark and Freishtat 2014) when his efforts are not recognised by students. This can be devasting for academics as it can stifle their enthusiasm and their willingness to perform their role: *“it affects your mental health, it affects your efficacy or motivation”* (Beth, female, Chinese). Although overall my participants were able to manage their feelings of anxiety as they gained more experience, dealing with negative and abusive comments continued to have a huge impact on their wellbeing. As a result, my informants would often seek support from their partners, peers or even resort to therapy in order to deal with the stress since they were not supported by their institutions (Heffernan 2023).

**The impact of (biased) student evaluations on academics’ career progression**

My participants recognised that student evaluations were closely tied with their career progression (Uttl and Smibert 2017). My analysis shows that achieving positive scores in student evaluations is important for everyone who has a student-facing role in Higher Education.

 Achieving high scores in student evaluations is the ultimate goal for my informants because these scores are used to assess if someone has achieved their probation requirements or when assessing an academic’s performance annually during Professional Development Review (PDR) meetings (which is standard practice in U.K. Higher Education). The target for the majority of my informants is to receive scores at around 4.0 in a Likert scale ranging from 1.0 (lowest) to 5.0 (highest); *“It’s a directive from the institution, across the sector that [..], you have to achieve a minimum of 3.5 and above to be considered decent”* (Mark, male, Arab). Jess explains:

*“The end of your official evaluation which our employers use to assess us annually. That number, that 4. whatever, I need that for my [PDR] assessment. I would rather that they are generous with that number, because that number actually is just for my employer.”* (Jess, female, Indian mixed)

According to Jess, line managers, Directors of Education, etc. usually focus on the quantitative score in order to assess if academics are reaching their targets in terms of student satisfaction; *“it’s the quantitative scores that they [universities] believe is the most important thing”* (Youssif, male, Arab). As a result, achieving these high scores creates a lot of pressure for academics. Faiz explains:

*“The score will also go to your PDR. So, there is always a kind of pressure to do well in module evolutions, otherwise it will affect your appraisal. [..] So, you will feel always uncomfortable [..] because your line manager will also look at that. [..] you might need to defend yourself.”* (Faiz, male, Bangladeshi)

According to Faiz, knowing that the results of student evaluations will be shared with his line manager made him uncomfortable. Faiz has felt particularly stressed in the past because he is aware that if you don’t achieve the necessary scores then you are put in a position to defend yourself which can cause a lot of trauma, especially for ECRs (Heffernan 2023). These pressures are exacerbated for marginalised academics for whom achieving probation requirements has visa implications: *“As an ECR that was super stressful. Because you just don’t know what real implications this would have on your probation case, and your visa”* (George, male, black).

My analysis shows that those academics who perform well in student evaluations are rewarded by institutions. For example, having positive evaluations helps academics in the job market because it shows to prospective employers that they can perform well as module leaders. Emre discusses his experience:

*“I feel like they helped me a lot because, I feel like one of the reasons why I got this job is because I am perceived quite positively by the students. [..] and definitely, a lot in terms of getting this position in a very short amount of time.”* (Emre, male, white)

Emre credits much of his success at this (early) stage of his career in being perceived positively by students and having that reflected in student evaluations. Receiving these positive scores and feedback has enabled him to pursue opportunities and to secure a permanent contract as Assistant Professor in the U.K. Higher Education context.

Universities also strategically celebrate staff that achieve high scores by giving them teaching awards; *“People are probably more impressed by my teaching award because of the implicit assumption that if you’ve been given an award, it means that your score was high relative to your peer group”* (George, male, black). George’s quote shows that teaching awards are used to signal academics’ teaching credentials. Academics then are able to use the cultural capital they gain from these institutionalised processes in order to support their probation or promotion cases, or when applying for jobs in different institutions; *“When I’ve had positive feedback it has been recognized, and particularly by the institution I’m at now. I was promoted”* (Kelly, female, black mixed).

**Navigating institutional pressures**

My analysis illustrates the impact of institutional pressures that academics encounter within the U.K. Business School context; *“I want the positive evaluations. I think institutionally we don’t want complaints, so we take a preventive approach”* (Beth, female, Chinese). According to Beth, academics tend to adopt a preventative approach when interacting with students in order to avoid negative evaluations. Beth accredits this mentality, that she has internalised, to institutional pressures. Natalie (female, other non-white) explains that this steams from the fact that when there is an issue *“they [universities] don’t see it as an institutional problem, they see it as your problem”*. Mary explains:

*“There is a pressure to make students happy. Obviously, I think because we are in the U.K., they almost treat the students as customers, and I think it will be more than this, as the years go by, because of Brexit, and because the students come from different backgrounds, they might need more help from you. But personally, there is a pressure. Yeah, I feel like I want to do well in student evaluations, and that’s why I used to feel nervous about it.”* (Mary, female, Indian)

Mary argues that despite all your other efforts and achievements as an academic, being able to make students happy is still a requirement. Mary believes that these institutional pressures are linked to the neoliberal logics and overcommercialisation of Higher Education that have transformed U.K. Business Schools into corporation-like entities in which students are seen as customers (Fleming 2019). As a result, there is a lot of added pressure for academics to perform well in metrics while there is usually little to no support from institutions. This often compromises the quality of teaching; *“They try to lower standards to make students happy”* (Jack, male, white). Faiz shares a similar opinion:

*“Universities function like a corporation. So, when they focus more on students, they forget about us [academics]. [..] [They] destroy the quality of education in the U.K., we are trying to compromise the quality, because we want good feedback from students and students know that, and they are trying to utilize it”*. (Faiz, male, Bangladeshi)

Faiz explains how this customer-oriented logic and hyper-fixation to receive positive evaluations is detrimental to the quality of learning and ultimately to the overall student experience. Faiz argues that students are aware of the importance of metrics (Suárez Monzón et al. 2022) and are even strategically rating modules to influence decision-making. Mark discusses the impact this has had on students’ expectations from their teachers:

*“We almost have to be like Houdini and teach like magicians. So I feel like that’s almost like an extra pressure on us as academics to try really and conform to institutional norms, because in the back of our minds there’s this lingering thought that says well, okay if I don’t start juggling, and pulling the rabbit out of the hat and doing magic, then students are not going to be impressed at the at the end of the module and so that’s going to reflect poorly on me. And then because it might reflect poorly in the evaluation, then I’m going to get a knock on the door and say, well, why haven’t you met the institutional module evaluation benchmark. [..] it puts you inside a box, and because you’re in a box you’re operating within tightly defined boundaries”*. (Mark, male, Arab)

According to Mark the institutional pressure to achieve high scores in student evaluations is driving academics to focus more on entertaining students during lectures instead of focusing on whether or not they are learning; *“No lecturer wants a negative evaluation [..] then maybe all you need to do is then just give them all that they want, so they can say you are good”* (Peter, male, African). This performative approach to teaching delivery is rewarded in student evaluations and as such “teaching can become a popularity contest, a theatre of entertainment rather than education” (Fleming 2019, 1307). Olivia shares her experience:

*“There is some pressure on how to approach teaching and interacting with students. Every meeting it’s all about the student experience, value, journey, as if we don’t know about that. Over focusing on these things adds pressure where there is already a lot. So, I think over stressing about having good reviews, or satisfying the student body and things like that it just adds more to it”*. (Olivia, female, white)

Olivia’s narrative illustrates how this added pressure on academics is communicated through different channels, such as during meetings. Olivia believes that these pressures to satisfy students clashes with the role of the educator and as a result it causes mental health problems for academics (Heffernan 2021; 2023).

**DISCUSSION**

The aim of this paper was to discuss the impact of biased student evaluations on academics’ wellbeing and career progression. I also showed the impact of institutional pressures to keep students happy and to perform well in student satisfaction metrics. My findings show that academics experience a lot of anxiety and stress in relation to student evaluations (Heffernan 2021; 2023). In line with prior studies, my analysis shows that often students do not focus on academics’ performance or on their overall experience with the module, but they rather use evaluations to criticise their teachers (Marsh 2007; Tucker 2014). As a result, I find that it is imperative to acknowledge how the combinations of intersectional bias around gender, race, class, age, (dis-)ability, sexuality, appearance, etc. influence how students perceive academics in order to be able to identify who is more likely to be disadvantaged by student evaluations and also how institutions can support and care for their staff accordingly (CRIS Collective 2023; Parsons et al. 2022). These intersecting factors gain different meanings in relation to each other (Collins and Bilge 2016) and produce different experiences for academics. My findings show that these factors inform students’ perceptions of academics’ identity and shape their expectations and evaluations (Boring 2017; Gelber et al. 2022; Sprague and Massoni 2005). Following an intersectional approach has also enabled me to illustrate how marginalised academics are disproportionately disadvantage by student evaluations in their career trajectory because they are more likely to receive negative evaluations (Chávez and Mitchell 2020; Heffernan 2021; 2023; Utoft and Pradhan 2023). In line with prior studies, my analysis shows that student evaluations influence academics’ career progression (Boring et al. 2016; Fan et al. 2019; MacNell et al. 2015; Mengel et al. 2017; Uttl and Smibert 2017). My findings also show that institutions use these metrics in order to discipline or reward those who align with normative expectations and perform well in evaluations through awards, fast-track promotions, etc.

My research offers three contributions. First, my primary contribution is to prior studies focusing on the role of biased student evaluations. The majority of the literature is mostly based on analyses of student feedback to illustrate the existence of biases (e.g., Boring 2017; Boring et al. 2016; Fan et al. 2019; Joye and Wilson 2015), whereas my study focuses on academics’ lived experiences with biased student evaluations. Specifically, I make an empirical contribution to the critical literature on student evaluations by drawing on intersectionality (Collins 2015; Collins and Bilge 2016; Crenshaw 1991) to show the impact of biased student evaluations on academics’ wellbeing and mental health (Heffernan 2021; 2023; Mengel et al. 2017; Tucker 2014; Stark and Freishtat 2014). Apart from a few exceptions (e.g., Heffernan 2023), my study is one of the few empirical investigations that show how academics from different backgrounds are impacted on a personal level from engaging in the student evaluation process in management education. My findings illustrate the immense pressure to perform well in evaluations and the resulting trauma of being subjected to derogatory and borderline abusive comments from students from the point of view of academics.

My study also offers a second theoretical contribution to the literature by showing that gender, race, (dis-)ability, sexuality, age, appearance etc. intersect to produce unique social locations of (dis-)advantage for academics (Collins and Bilge 2016). The majority of the literature has primarily focused on how gender bias impacts student evaluations (e.g., Boring 2017; Mengel et al. 2017; Joye and Wilson 2015; MacNell et al. 2015). My findings, in line with nascent research that goes beyond gender (Chávez and Mitchell 2020; Fan et al. 2019; Heffernan 2021; 2023; Utoft and Pradhan 2023), illustrate that sexism, racism, ageism, classism, ableism, homophobia, etc. gain different meanings in relation to each other and produce different positionalities and experiences (Pirani and Daskalopoulou 2022) for academics.

My research also adds to the existing literature that argues for the removal of student evaluations due to their undeniable biased nature and also due to their detrimental effects on academics’ wellbeing (Heffernan 2021; 2023; Santisteban and Egues 2022). My findings provide an account of the pressures that academics experience in relation to their teaching performance. As suggested by Fleming (2019), teaching has taken on a performative turn where fulfilling students’ (biased) expectations and putting on a show is more valued than actually prioritising student learning. Universities are complicit in perpetuating these biases as long as students continue to be able to direct abuse towards (marginalised) academics and also by rewarding staff that perform well in a biased system.

My study is not without limitations. My research is based on the U.K. Higher Education context, and more specifically Business Schools, where a lot of emphasis is put on student evaluations. Future research should aim to explore the experiences of academics in different Higher Education contexts where other metrics could also play a role such as websites (e.g., Rate my Professor). Future research should also explore students’ accounts of providing negative evaluations. Analyses of students’ motivations and level of awareness around evaluations would be extremely useful in order to better understand how the learning experience can be evaluated. Universities are currently enabling a system that further marginalises its already marginalised academics (Heffernan 2023) by using student evaluations as a teaching quality assurance tool in management education. Hopefully, my research along with existing critical literature can be used as evidence of the impact of biases on academics’ lived experiences in order to make long overdue changes in our institutions to better support staff and students.

Finally, my findings have practical implications for management education. I encourage Business Schools to identify and document in a more systematic manner which biases are particular to their respective fields and teaching approaches, and also who is more likely to be affected negatively by student evaluations. This proactive approach will enable line managers to offer relevant support and to avoid further disadvantaging marginalised academics. Moreover, Business Schools need to offer tailored implicit bias training to academics that participate in committees that advice on promotions, PDRs, etc. and also academics that sit on interview panels. This will enable those in positions of power to understand that student satisfaction metrics are biased and not a true reflection of an academic’s ability to perform well in their role (Peterson et al. 2019). Finally, universities can take on a more radical approach and discontinue the use of student evaluation surveys. For example, the University of Southampton (2023) has recently stopped using end of module surveys because “of the concerns regarding conscious and unconscious bias”. Instead, Business Schools can opt for collecting mid-semester, qualitative module feedback run by module leaders on a module-by-module basis. This process will enable academics to gain an understanding of how teaching is going and will also allow them to implement changes for the remaining part of the teaching period (Yorke 2003). Although, this approach cannot ensure that evaluations will be less biased, it will at least alleviate some of the pressure for academics because they will not be collecting any quantitative scores to be used as performance indicators.

To conclude, biased student evaluations risk diminishing the potential of management education because the process focuses on the academic instead of the overall student experience with the module. It is vital to gain a deeper, holistic understanding on this topic in order to create appropriate support mechanisms for academics and students and hopefully this study makes a meaningful contribution towards this goal.

**REFERENCES**

Addison, M., Breeze, M., and Taylor, Y. 2022. *The Palgrave Handbook of Imposter Syndrome in Higher Education*. Switzerland: Palgave.

Basow, S.A. 1995. “Student evaluations of college professors: When gender matters.” *Journal of Educational Psychology*, 87(4): 656-665.

Boring, A. 2017. “Gender Biases in Student Evaluations of Teaching.” *Journal of Public Economics,* 145: 27–41.

Boring, A., Ottoboni, K., and Stark P. 2016. “Student Evaluations of Teaching (Mostly) Do Not Measure Teaching Effectiveness.” *Science Open Research*. DOI: 10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1

Braun V., and Clarke, V. 2022. *Thematic analysis: A practical guide*. Thousand Oaks, California: SAGE Publications Ltd.

Charmaz, K. 2006. *Constructing grounded theory: A practical guide through qualitative research*. London: Sage Publications.

Chávez, K., and Mitchell, K.M. 2020. “Exploring bias in student evaluations: Gender, race, and ethnicity.” *PS: Political Science & Politics*, 53(2): 270-274.

CRIS Collective (2023). “Let there be a “We”: Introducing an ethics of collective academic care.” *European Journal of Marketing*, 57(10): 2838-2859.

Collins, P.H. 2015. “Intersectionality's definitional dilemmas.” *Annual Review of Sociology*, 41: 1-20.

Collins P.H., and Bilge, S. 2016. *Intersectionality*. Cambridge: Polity.

Constantinou, C., and Wijnen-Meijer, M. 2022. “Student evaluations of teaching and the development of a comprehensive measure of teaching effectiveness for medical schools.” *BMC Medical Education*, 22(1): 113.

Crenshaw, K. 1991. “Mapping the margins: Identity politics, intersectionality, and violence against women.” *Stanford Law Review*, 43(6): 1241-1299.

Cunningham-Nelson, S., Baktashmotlagh, M., and Boles, W. 2019. “Visualizing Student Opinion through Text Analysis.” *IEEE Transactions on Education,* 62(4): 305–311.

Denzin, N., and Licoln, G. 2005. *Handbook of qualitative research*. London: Sage Publications.

Gelber, K., Brennan, K., Duriesmith, D., and Fenton, E. 2022. “Gendered mundanities: Gender bias in student evaluations of teaching in political science.” *Australian Journal of Political Science*, 57(2): 199-220.

Fan, Y., Shepherd, L., Slavich, D., Waters, D., Stone, M., Abel, R., and Johnston, E. 2019. “Gender and Cultural Bias in Student Evaluations: Why Representation Matters.” *Plos ONE,* 14(2): e0209749.

Feldman, K.A. 1992. “College students' views of male and female college teachers: Part I – Evidence from the social laboratory and experiments.” *Research in Higher Education*, 33: 317-375.

Feldman, K.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: 151-211.

Fleming, P. 2019. “Dark academia: Despair in the neoliberal business school.” *Journal of Management Studies*, 57(6): 1305-1311.

Heffernan, T. 2022. “Sexism, racism, prejudice, and bias: A literature review and synthesis of research surrounding student evaluations of courses and teaching.” *Assessment and Evaluation in Higher Education*, 47(1): 144-154.

Heffernan, T. 2023. “Abusive comments in student evaluations of courses and teaching: The attacks women and marginalised academics endure.” *Higher Education*, 85(1): 225-239.

Hughes, E., and Donnelly, R. 2023. “Business schools and faculty experiences of sexism: Gender structure tensions within and outside these schools.” *Gender, Work & Organization*, 1-20.

Johansson, M., and Śliwa, M. 2014. “Gender, foreignness and academia: An intersectional analysis of the experiences of foreign women academics in UK business schools.” *Gender, Work and Organization*, 21(1): 18-36.

Johnson, V.E. 2003. *Grade inflation: A crisis in college education*. New York: Springer-Verlag.

Joye, S.W., and Wilson, J.H. 2015. “Professor Age and Gender Affect Student Perceptions and Grades.” *Journal of the Scholarship of Teaching and Learning,* 15(4): 126-138.

MacNell, L., Driscoll, A., and Hunt, A. 2015. “What’s in a Name: Exposing Gender Bias in Student Ratings of Teaching.” *Innovative Higher Education,* 40(4): 291–303.

Marsh, H.W. 2007. Students’ evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R.P. Perry, and J.C. Smart (Eds), *The scholarship of teaching and learning in higher education: An evidence-based perspective*: 319-383. Dordrecht: Springer.

McCracken, G. 1988. *The long interview*. London: Sage Publications.

Mengel, F., Sauermann, J., and Zölitz, U. 2017. “Gender Bias in Teaching Evaluations.” *IZA Discussion Papers*. Institute of Labor Economics (IZA), Bonn.

Parsons, E, Pirani, D, Ashman, R, Daskalopoulou, A., Kerrane, K., and McGouran, C. 2022. Manifesting feminist marketing futures: undertaking a ‘visionary’ inventory. In P. Maclaran, L. Stevens, and O. Kravets (Eds), *The Routledge Companion to Marketing and Feminism*: 460-476. London, UK: Routledge.

Patton, M.Q. 2003. *Qualitative research and evaluation methods*. London: Sage Publications.

Peterson, D.A., Biederman, L.A., Andersen, D., Ditonto, T.M., and Roe, K. 2019. “Mitigating gender bias in student evaluations of teaching.” *PloS one*, 14(5): e0216241.

Pirani, D., and Daskalopoulou, A. 2022. “The queer manifesto: Imagining new possibilities and futures for marketing and consumer research.” *Marketing Theory*, 22(2): 293-308.

Santisteban, L., and Egues, A.L. 2022. “How student evaluations of teaching contribute to hindrance of faculty diversity?”. *Teaching and Learning in Nursing*, 17(4): 455-459.

Sebastian, R.J., and Bristow, D. 2008. “Formal or informal? The impact of style of dress and forms of address on business students' perceptions of professors.” *Journal of Education for Business,* 83(4): 196-201.

Sprague, J., and Massoni, K. 2005. “Student Evaluations and Gendered Expectations: What We Can’t Count Can Hurt Us.” *Sex Roles,* 53(11/12): 779-793.

Stark, P., and Freishtat, R. 2014. “An Evaluation of Course Evaluations.” *Science Open Research*. DOI: 10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1.

Suárez Monzón, N., Gómez Suárez, V., and Lara Paredes, D.G. 2022. “Is my opinion important in evaluating lecturers? Students’ perceptions of student evaluations of teaching (SET) and their relationship to SET scores.” *Educational Research and Evaluation*, 27(1-2): 117-140.

Tucker B. 2014. “Student Evaluation Surveys: Anonymous Comments That Offend or are Unprofessional.” *Higher Education,* 68(3): 347–358.

University of Southampton. 2023. Module Feedback, Principles for Module Feedback. Accessed 23 June 2023. Available: <https://www.southampton.ac.uk/quality/student_engagement/modulefeedback.page>.

Uttl, B., and Smibert, D. 2017. “Student Evaluations of Teaching: Teaching Quantitative Courses Can Be Hazardous to One’s Career.” *Peer Journal* 5: e3299.

Utoft, E.H., and Pradhan, A.A. 2023. “Lecturers as robots: Student evaluations of teaching as dehumanizing, neoliberal technology.” Accessed 31 October 2023. Available: <https://www.researchgate.net/publication/374421078_Lecturers_as_robots_Student_evaluations_of_teaching_as_dehumanizing_neoliberal_technology>.

Wright, S.L., and Jenkins-Guarnieri, M.A. 2012. “Student evaluations of teaching: Combining the meta-analyses and demonstrating further evidence for effective use.” *Assessment & Evaluation in Higher Education*, 37(6): 683-699.

Yorke, M. 2003. “Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice.” *Higher Education*, 45(4): 477-501.