Check for updates

'Professor Moms' & 'Hidden Service' in Pandemic Times: Students Report Women Faculty more Supportive & Accommodating amid U.S. COVID Crisis Onset

Danielle Docka-Filipek[1] [ORCID] · **Crissa Draper[2]** · **Janice Snow[3]** · **Lindsey B. Stone[4]**

Abstract

Emerging data suggests the COVID-19 crisis exacerbated preexisting, long-documented gender inequities among U.S. faculty in higher education. During the initial Spring 2020 'lockdown' in the U.S., 80 students conveyed their experiences with faculty across 362 courses. We evaluated whether students' reports of faculty supportiveness, accommodations granted, and pandemic-impacted, anticipated grade outcomes differed according to faculty gender via mixed linear models (data on 362 courses were nested within 80 student reporters). Students perceived their women instructors as more supportive, accommodating, and anticipated lesser course grade decreases across the semester than in courses taught by men. Accordingly, we interpret that amidst the 'lockdown' crisis, women faculty earned higher perceived supportiveness and positive student outcomes than their male counterparts. Further, the data likely reflects women faculty's greater conscription into demonstrated care work, despite the coding of such labor as "feminine," thereby rendering such work devalued. To reframe, to the degree that students expect more 'intensive pedagogies,' which invites faculty and administrators to gender disparate demands, such pressures likely translate to 'hidden service' burdens, and correspondingly, less time for career-advancing activities (such as research). Broader implications are discussed, alongside women faculty's documented experiences of acceleration in career and work/family pressures in pandemic-times, which combine to exacerbate long-standing, yet now-amplified penalties, potentially driving a widening, gendered chasm in academic career outcomes. We conclude by offering constructive suggestions to mitigate any discriminatory impacts imposed by students' gendered assessment inputs and expectations.

Keywords Gender inequality · Hidden service · Women faculty · COVID-19 pandemic · Teaching evaluations

✉ Danielle Docka-Filipek
dock0069@d.umn.edu

Extended author information available on the last page of the article

The literature on "hidden service," "secret service," "invisible academic work," or "academic housework" typically defines associated, relational academic labor as temporally and emotionally demanding carework that receives little visibility or tangible reward for career advancement, yet is nevertheless expected—to a disproportionate degree—of women and racialized faculty (Domingo et al., 2022; Górska et al., 2021; Hanasono et al., 2019). Such work sits in contrast to more "task-oriented" service labor that lends itself more readily to quantifiable metrics, 'deliverable' products, measurable institutional rewards, neoliberal/racialized/gendered logics of "excellence," (Mickey et al., 2022) which are altogether more often coded "male" in academic environments [in line with Acker's (Acker, 1990) white, male, able-bodied, and heterosexual "ideal worker;" see Johnson (2022) on the increasing salience of those existing templates amid the COVID crisis]. Some characteristic examples of 'low-profile,' yet vitally important, "hidden service" work in institutions of higher education include close, detail-oriented, and active mentoring in support of students' positive academic experiences; performing emotion work for university constituents at all levels; efforts explicitly intended to address ("chilly") organizational climate; and serving on 'lesser'/marginalized/underresourced institutional committees (Hanasono et al., 2019). Typically, the caregiving tasks embedded in teaching, mentoring, and the general emotion work expected or required on-the-job from women faculty members are all tasks that require a high degree of pastoral care and emotional presence, yet they also divert time and energy away from more prestigious or highly regarded work activities such as research, writing, and publishing (Ashencaen Crabtree & Shiel, 2019; Bagilhole & Goode, 1998). What's more, in the burgeoning literature that tracks the amplification of such preexisting gendered/racialized burdens in the academic workplace, many analysts are finding that institutions' expressed needs and corresponding expectations for such gendered, "hidden" caregiving labors have persisted and heightened amid the escalation of pandemic-associated troubles with recruitment, retention, academic performance/engagement, and mental health crises (particularly for students, but also for faculty and support staff; see Cate et al., 2022; Docka-Filipek & Stone, 2021; Górska et al., 2021; Plotnikof & Utoft, 2022). In many cases, institutionalized academic performance and evaluation metrics have failed to keep pace with measuring the caregiving labor that may well amount to the very 'glue' holding our institutions of higher learning together, particularly amid such ongoing, multifaceted crises (Benozzo et al., 2022; Branicki, 2020; Mickey et al., 2022; Özkazanç-Pan & Pullen, 2020; Pereira, 2021).

We sought to examine how student expectations for 'extra,' low-reward or no-reward pedagogical labor (measured via student reports of supportiveness, accommodations granted, and anticipated grade drops) varied on the basis of instructor's gender during one very specific, historically-contingent moment of acute crisis: the start of the COVID-19 pandemic in the U.S. In such times of institutional crisis, we argue, otherwise unspoken or less obvious normative conventions for the lop-sided, gendered division of academic labor may be rendered more visible or evident, and needs often threaten to outstrip resources in such scenarios. In other words, we reasoned, those members of the 'university family' who find themselves 'leaned on' during difficult times may well find the otherwise persistent 'hum' of the pressure to disproportionately invest such caregiving

labor amplified to a deafening degree, rendering such conventions more readily distinctive or manifest.

We tested our assumptions regarding women faculty's heightened expectations for pedagogical "hidden service" by asking students to report on their experiences with faculty in their enrolled courses during the initial spring 2020 'lockdowns.' Ultimately, amid the campus evacuations during the first wave of the COVID crisis in the U.S. (during spring of 2019), students perceived their women instructors as more supportive, accommodating, and more flexibly tailored in their evaluative standards—all during a moment of acute crisis, upheaval, and uncertainty. We argue such findings point to both qualitative and quantitative gender disparities in students' expectations for teaching-related faculty labors. More specifically, students anticipated their women instructors would engage their learning process through pedagogical efforts that are both more temporally and more emotionally taxing. Such greater demands on women faculty have undoubtedly translated to a 'hidden service' burden, and by extension, lesser time for activities with greater career advancement rewards (as established in findings from pre-pandemic times, delineated above, as well). In the context of the pandemic, while students exhibit heightened expectations for the labor of women faculty, such pressures may then combine with the gender disparate career and work/family pressures incurred by the pandemic, alongside the now-established gendered punishments of student evaluations of teaching or "SETs" (Heffernan, 2022; Kreitzer and Sweet-Cushman (2021). Taken together, the impact of such "hidden service" may well contribute to widening a potentially growing chasm in academic career outcomes, on the basis of gender.

Further, although we concede that the onset of the global health crisis precipitated by the COVID pandemic absolutely posed a unique, unprecedented, and historically-specific moment, we *also* argue that the state of US higher education has long been represented by governing authorities as tumbling into a state of ever-deepening fiscal and normative crises, which are now brought into especially stark relief in pandemic times. Such catastrophic depictions have been repeatedly, and now predictably, characterized by broader challenges to the ideological hegemony of higher learning as universal public and individual good (Fingerhut, 2017; Newport & Busteed, 2017). Other threats to the sustainability of U.S. higher learning include declining enrollment numbers due to overall decreases in the college-aged population; ever-accelerating losses of state revenue allocations; the politically fickle 'charity' contributions of the philanthropic class; cyclical economic crashes endemic to late capitalism; the associated, widespread anxieties of multiple 'stakeholders' regarding the solvency of institutional endowments; and the inherent threats to democratic ideals posed by narrowing neoliberal rationality (Brown, 2015; Geiger, 2010). Against this broad national backdrop, the COVID crisis has only served to amplify survivalist fears that plug into already-existing institutional narratives, policies, and practices that revolve around the doxa of austere measures and the corresponding implementations of 'fat-trimming' institutional transformations (Giroux, 2014)—all of which threaten to pry cracks into crevices along the 'leaky pipeline' of faculty career advancement. Indeed, while the literature has long tied the above-mentioned constraints to the unique threats that women, mothers, and faculty of color face to their academic careers and well-being, accumulating empirical evidence increasingly

demonstrates an acceleration of prexisting, gender-disproportionate burdens and obstacles to career progress, amidst the COVID-inspired shifts in the academic workplace (Davis et al., 2022; Górska et al., 2021; Johnson, 2022; Kasymova et al., 2021).

Therefore, while COVID is certainly 'new,' the worries, implemented practices, and most importantly for our analysis, the resulting gender disparities in faculty careers that the pandemic's moment of crisis set into motion are altogether far from novel. We therefore argue the onset of COVID-based institutional scramblings and their demographically varying impacts on different segments of the U.S. professoriate are worth considering and perhaps extending to inevitable future turning points inspired by calamity, as it seems unlikely that the threats to U.S. higher education (delineated above) are likely to evaporate from the national landscape anytime soon. In other words, while COVID may have been unprecedented, the realities it set into motion in higher education most certainly are not, which means pandemic-inspired conditions present important, empirical learning opportunities.

Gender Bias & Inequities Permeate Western Culture and the U.S. Academy

Implicit bias (Devine, 1989) refers to the precognitive or preconscious assumptions social actors make that impact perceptions of and behavior towards others, which may result in discriminatory outcomes. In the case of gender, we tend to associate femininity with particular traits that in many circumstances (especially those related to wage-earning/prestigious work) are presumed to have lesser value than those traits assumed to be linked to masculinity (Cheryan & Markus, 2020). Such 'feminized' traits may either be presumed to be linked to particular elements of the physiological, hormonal, or reproductive consequences and paramount 'being' of people assigned 'female' at birth (irrespective of the reality that many such generalizations do not apply to a good many 'women'), via ideological conventions typically referred to as "gender essentialism," or alternately, "biological determinism" (DiQuinzio, 1993; Grosz, 1995).

Nevertheless, developing institutional and individual awareness regarding the reality of the operations of implicit bias is not enough to avoid inflicting future bias (Kim, 2003). Bem's (1993) pioneering work on the "lenses of gender" posits that sexism is so deeply embedded in our social institutions, our cultural constructions, and our psychological makeup that androcentrism, gender polarization/complementarity (i.e., the belief that masculinity and femininity are not only opposite 'poles,' but also serve to 'complete' one another), and gender essentialism are now keystone components of our individual and collective experiences of gender. Butler's (1990) landmark works make clear these well-established, gendered 'scripts' permeate multiple realms of human social life, and correspondingly elicit punishing social evaluative consequences for gender 'deviance' (or "gender trouble").

Such insights combine to dictate that even if individuals are made aware of how gender bias functions, implicit learning nevertheless impacts their quantitative assessment of the work performance of individuals, in the absence of explicitly

corrective practices (Key & Ardoin, 2019). Unsurprisingly, extensive evidence of gender bias extends to the academic workplace (Régner et al., 2019). More specifically, role congruity/role conflict theories (Eagly & Karau, 2002) posit women leaders will be evaluated more harshly when their leadership behaviors are seen in fundamental conflict with the prescribed behaviors and personality traits typically associated with femininity. In other words, behaviors that conform to gendered expectations for one's actions—women behaving as non-threatening, passive, emotionally warm, caring, comforting, accommodating, flexible, and nurturing—are rewarded, whereas deviations from gendered prescriptions for behavior (e.g., brief, abrupt, critical, exacting, assertive, inflexible/rigid, non-accommodating) are punished (Eagly & Karau, 2002).

Further, prior to the COVID crisis, women faculty spent measurably greater amounts of time on service per week (Guarino & Borden, 2017), and what's more, they spent substantially more time in less institutionally-rewarded, or 'altruistic' forms of service such as mentoring and department-level service to students, to the measurable detriment of their career advancement (Hanasono et al., 2019). Male faculty, again, tend to spend greater amounts of time in more visible and measurably 'valuable' forms of service work (Misra et al., 2012). In experiments, women faculty more frequently volunteer, register, and accept requests for work tasks with lesser rewards, presumably resulting in more "office housework" for faculty women than faculty men (Babcock et al., 2017). Women faculty also receive and honor more requests for 'special favors' and emotional labor from students, and they do so more often for more 'academically entitled' charges (El-Alayli et al., 2018). Additionally, women faculty attribute their greater time spent in service due to a felt sense of 'duty' to the collective enterprise of higher learning (Misra et al., 2012). To add insult to injury, such work reportedly leaves many women instructors feeling disappointed or disillusioned, particularly when coupled with other gendered on-the-job insults (Acker & Feuerverger, 1996). Indeed, when such labors are not rewarded, and are instead explicitly devalued by administrators, senior colleagues, and performance review bodies, they may also be imbued with meaning that draws on gendered and/or racialized stereotypes (Domingo et al., 2022). Ironically, again, while such "secret service" may be associated with lesser institutional reward, such labors are arguably of central important in cementing the 'ties that bind' in times of crisis or uncertainty, as during the onset of the global COVID crisis.

COVID-19 and Gender Disparities at Work and in the Home

Amidst the COVID-19 pandemic, women faculty have increasingly found themselves conscripted into "mothering" labors both in the workplace and at home, and at considerable cost to their well-being and mental health (Docka-Filipek & Stone, 2021). We contextualize our findings that point to the salience of such caregiving labor during the "lockdown" semester in the burgeoning literature that documents the acceleration of such demands across the 'public' and 'private' spheres, in order to paint a comprehensive picture of the manifold obstacles to the career progress of women faculty. Below, we review multiple, converging sources of pressure to

mother 'around the clock,' which together contribute to the pandemic's 'blurring' impact on the roles women faculty occupy both at home and on the job, rendering a healthy "work/life balance" and timely career progress ever-elusive.

While there is ample evidence that women experience higher rates of anxiety and depression, research has further suggested that these mental health vulnerabilities have been amplified by pandemic conditions (Thibaut & Van Wijngaarden-Cremers, 2020). Relatedly, previously documented disparities in hetero-coupled husbands and wives' time spent in homecare have deepened amid the COVID crisis (Del Boca et al., 2020). During the initial lockdown women took on greater childcare duties (Zamarro et al., 2020); including greater time spent on home-schooling activities—despite men's parallel perception that such duties were shared equally (Miller, 2020). Women's intensified childcare roles often came at the cost of mothers' emotional well-being (Calarco et al., 2021) and resulted in gender discrepancies in satisfaction with the home/work environment among heterosexual academic couples (CohenMiller & Izekenova, 2022; Yildirim & Eslen-Ziya, 2021), as well as an increase in domestic interpersonal conflicts (Calarco et al., 2020).

Additionally, women's assumption of new caretaking roles at home co-occurred with macro-economic shifts that dictated staggeringly disproportionate job losses for women (particularly, women of color and single mothers) in the wake of the pandemic (U. S. Bureau of Labor Statistics, 2021; Ewing-Nelson, 2021; Petts et al., 2021), suggesting many mothers were 'pushed out' of the paid labor force. Indeed, women reduced their paid work hours at nearly five times the rate of men in heterosexual couples, even when telework was possible (Collins et al., 2021). Regarding quantifiable faculty career impacts, pre-pandemic gender differences in research productivity are also now exacerbated (Huang et al., 2020). Women's publication rates have stagnated while men's submission of articles to publication venues has accelerated (Cui et al., 2022; Squazzoni et al., 2020). Put simply, the time, space, and emotional reserves women had for their traditionally anticipated caretaker role with their students was profoundly impacted via multiple sources of accelerating, competing demands both at home and at work that together measurably increased amid the coronavirus pandemic.

Despite any potential sense of reward that may accrue from performing workplace 'mothering,' being constantly "on the clock" for such tasks not only detracts from tasks that carry greater career/institutional rewards, such 'mothering' labors also exact considerable, gendered psychological costs (Gregg, 2011). Boncori (2020) aptly refers to the circumstances of pandemic times for mothers as converging in "the never-ending shift." Such porousness between work and home for women faculty was documented in pre-pandemic times as generating women's greater susceptibility to bidirectional and negative work/family "spillover" (Eddleston & Mulki, 2017), which has amplified during the COVID-era (Craig & Churchill, 2021). Nevertheless, we emphasize that any refusal of such caring labors likely carries real career penalties as well, as the tightrope we walk is a narrow one. Prominent vehicles for such punishments are "SETs," or student evaluations of teaching (Richards, 2019; Sinclair & Kunda, 2000), as well as the gender biased and compulsory caregiving expectations embedded in other formal and informal evaluative performance metrics (see Docka-Filipek & Stone, 2021).

Student Evaluations of Teaching ("SETs"): Gendered Instruments Exact Gendered Punishments

Overall, SETs tend to reward women faculty more highly when their classroom behaviors, policies, and practices are congruent with widespread, hegemonic workplace and educational expectations for feminine comportment in leadership positions. Conversely, SET's tend to punish with hostility women faculty whose gender performances are incongruent, unexpected, or 'nontraditional' via cultural standards for femininity (Sprague & Massoni, 2005). Therefore, a wide swath of analysts have concluded that SETs are themselves gendered instruments (Kreitzer & Sweet-Cushman, 2021). Recently, one group of analysts found that although women lecturers did not score lower on their SETs, gendered behaviors in line with stereotypical femininity elicited expectations from students that their instructor would be more approachable, students preferred to attend their courses, and students rated their feminine instructors as more "likeable" (Renström et al., 2021). Further, gender-transgressive behaviors are more readily forgiven by students in their ratings if women faculty are perceived as young/'attractive' by the standards of conventional femininity (Arbuckle & Williams, 2003). Indeed, prior research indicates that women faculty's expected adherence to behaviors consistent with prescribed gender norms has a greater impact on SETs than faculty gender alone (Basow & Silberg, 1987; Freeman, 1994).

More specifically, students' greater expectations for caregiving labors from women faculty are quite obviously emotionally and temporally taxing (El-Alayli et al., 2018), which serve to contribute to gender disparities in academic careers (Domingo et al., 2022; Hanasono et al., 2019). For example, students tend to either score women instructors more highly or respond less frequently with hostility, resistance, or retaliatory punishments in their evaluative ratings of teaching when the workload is lower and/or the grading scheme is more lenient (Sinclair & Kunda, 2000). Students are more readily accepting of criticism from male faculty without also perceiving associated shortcomings as also pointing to instructional deficits (Sinclair & Kunda, 2000). Thus, women faculty may receive lower evaluations for enacting similar grading scales as their male counterparts (Freeman, 1994). Further, students' evaluations of the quality of teaching tend to increase when women faculty are perceived as more caring, emotionally invested, nurturing, accommodating, or flexible (Ashencaen Crabtree & Shiel, 2019; Bagilhole & Goode, 1998; Sprague & Massoni, 2005). While it remains unclear to what extent gendered student expectations for faculty point to gendered discrepancies in students' *perceptions* of faculty behavior or students' accurate reading of *actual* gender differences in faculty behavior, researchers have identified both a link between higher frequencies of student demands for "special favors" (tasks outside of typical work duties) and greater self-reported emotional labor (El-Alayli et al., 2018), as well as male students' lesser likelihood of following instructions given by women instructors (Piatak & Mohr, 2019).

Despite the confounding impact of faculty gender and the accumulating findings regarding the lack of relationship between positive student assessments and

evidenced learning (Boring & Ottoboni, 2016; Esarey & Valdes, 2020; Kogan et al., 2022; Stroebe, 2020; Uttl et al., 2017), SETs continue to be heavily weighted in most institutional review metrics (Sprague & Massoni, 2005), irrespective of available, viable alternatives [see Miller & Seldin, 2014 on strategies such as peer observations; Centra, 2000 and Seldin et al., 2010 on review of teaching portfolios; and Chism & Chism, 2007 on internal or external reviews of course materials]. Notwithstanding these flaws, the documented gender bias in SETs is theorized to be a primary driving force behind gender and racial disparities in faculty job placements, career achievements (Shreffler et al., 2019; Weisshaar, 2017), and promotion/tenure/pay disparities (Murray, 1984; Wachtel, 1998)—especially so in typically male-dominated fields and at institutions generally considered prestigious or top-ranking (Huston, 2006; Pittman, 2010; Reid, 2010). Because SETs factor heavily in faculty performance metrics, they impact the composition of the U.S faculty and reduce the overall numbers of historically underrepresented identities in the professoriate, especially at higher ranks (Branch, 2017; West & Curtis, 2006). Such findings have been linked to women's disproportionate time and emotional energy spent on their teaching to reach the same or better outcomes as their male colleagues (Laube et al., 2007). Taken together, such findings point to greater workplace burdens via students' expectations for caregiving labor from women faculty, which we argue are likely enforced by the threat of negative student SETs and the reception of such scores in the bureaucratic evaluative process. In short, gender-specific elements of women instructors' teaching labor presumably comes at the cost of women faculty's achievements of other important career accomplishments necessary for tenure, promotion, and even retention (El-Alayli et al., 2018; Laube et al., 2007; O'Meara et al., 2017).

We contextualize our insights regarding the gendering of SET instruments within the broader literatures on gender disparities in academic careers (Branch, 2017; West & Curtis, 2006), which are partly fueled by the heightened burdens of "secret service" (Domingo et al., 2022; Guarino & Borden, 2017; Hanasono et al., 2019; Tuck, 2018), as well as gender asymmetries in labor performed in the home (U.S. Bureau of Labor Statistics, 2021; Yavorsky et al., 2015), and now, likely also driven by a gendered acceleration in labor demands driven by the pandemic (Craig & Churchill, 2021; Yildirim & Eslen-Ziya, 2021). The current study examines whether and *how* faculty gender affected students' perceptions of faculty support and academic achievement during the initial wave of the COVID pandemic.

The Current Study

In the current study, undergraduates at a large public institution reported on their experience in each course they were currently completing, their assigned instructional faculty's gender, their perceived course grades mid-semester (pre-COVID), and their estimated final grades. Overall, our study was designed to examine the impacts of instructor gender on student ratings of support, any gender disparities in the reported level/degree of individual accommodations granted by instructors, the impact of instructor's gender on anticipated grade drops occurring in the moments

between initial evacuations (April, 2020) and the end of the semester (May, 2020), and whether instructor's gender impacted students' reporting of any pandemic-related negative impacts on their final grades.

Expectations

Given students' gendered anticipation of greater levels of teaching labors from their women faculty that they view as supportive, nurturing, or understanding (Ashencaen Crabtree & Shiel, 2019; Bagilhole & Goode, 1998; Sprague & Massoni, 2005), we expected that students would describe women faculty as more supportive. Further, because the literature points to more frequent student demands/requests for "special favors" (El-Alayli et al., 2018), we expected students would assess their women faculty as more accommodating. Additionally, given students' documented sense of entitlement to greater evaluative leniency from women faculty (Sinclair & Kunda, 2000), we surmised students would rate their women instructors as more accommodating, and less penalizing, when compared to men faculty. Additionally, we anticipated students would self-report higher current (post-'lockdown,' pre-final) grades in courses taught by women faculty, also largely due to students' sense of entitlement to lesser penalties and rigid standards from their women instructors (Sinclair & Kunda, 2000),which we further predicted would lead to lower anticipated grade drops from self-reported midterm (pre-'lockdown') grades to self-reported current (post-'lockdown,' weeks prior to the end of the term) grades, as well as lower anticipated pandemic-related negative impacts on final grades from women faculty, in comparison with students' assessments of men faculty. Lastly, we also considered whether students' gender moderated the effect of perceived faculty support, accommodations, and grades, given the emerging data suggesting that students tend to evaluate the faculty that share their same gender more highly (Bachen et al., 1999; Young et al., 2009)—in other words, that male students tend to evaluate women faculty lower than their male faculty (Fan et al., 2019; Mengel et al., 2019), and conversely, women students rate women faculty higher (Centra, 2000).

Data & Methods

Participants and Procedure

Undergraduate students were recruited at a large public university in the Western United States, and were subsequently asked to complete a questionnaire designed to examine how the current COVID-19 pandemic was affecting them, and to gauge their assessment of the projected impact it may have on their personal academic outcomes. Inclusion criteria was 18 years of age or older, and any current enrollment in undergraduate courses at the university. Participants were sent to an online Qualtrics survey, which directed participants to an informed consent page. If they consented to participate then participants were asked to provide demographic data and complete self-report surveys on mental health, their current living situation. Finally,

they were asked to complete a series of questions for *each* course they were currently completing (including faculty gender, perceived grade, faculty support etc.). This research was approved by the Institutional Review Board of the university from which the data was collected, and informed consent was obtained from all participants. Recruitment started April 14th, 2020 and ended May 7th, 2020. Of the 89 visitors to 'consent' to participate in the study, 80 completed the survey. The average completion time was 30 minutes (median 23 minutes). Participants received course credit for completing the survey.

Demographic statistics of the 80 participants are displayed in Table 1. Participant ages ranged from 18 to 54, with 80% of students under age 25. The majority of participants were Caucasian, 91%. Furthermore, it bears mention that the university used for recruitment holds a regional and national reputation for social and political

Table 1 Sample Characteristics

Demographics	Range	Mean	*SD*
Age	18–54	22.70	6.51
Gender, *n* (%)			
Woman		55	68.8%
Man		24	30%
Race, *n* (%)			
White		73	91.3%
Black		1	<1%
Native Alaskan or American Indian		1	<1%
Asian		4	<1%
Multi-racial		4	<1%
Other		2	<1%
Ethnicity, *n* (%)			
Hispanic or Latino		9	11.3%
Years in school/academic Standing	1–5	2.23	1.33
Student Type, *n* (%)			
Traditional		28	35%
Distance learning		10	12.5%
Traditional and has taken at least one online class		42	52.5%
Married, *n* (%)			
Single		40	50%
Married or living with partner		25	31.3%
In a relationship		15	18.8%
Employment (Feb), *n* (%)			
Full-time		9	11.3%
Part-time		40	50%
Unemployed		31	38.8%
Parent status, *n* (%)			
with children		11	13.8%

conservatism, though the school's reputation likely exceeded its reality, as the frequency of "conservative" identification among students was around 15% higher than the national average for college students (results should be interpreted accordingly, and future analyses may explore potential links between social/political ideologies and students' gendered instructional/educational expectations).

Students' Assessment of Faculty Support

For each course students were currently completing, they were asked to identify their instructors' gender (Male, Female, other). The choice to assess perceived gender via sex terminology was intentional. With the conservative nature of the student population we were not confident with their familiarity with common gender terms (e.g., cisgender). To assess students' perception of faculty support they were asked to rate for each instructor "to what degree do you feel supported by your instructor?". Ratings were given on a 5-point Likert scale ranging from 'a great deal' (1) to not at all (5). Answers were then reversed scored so that higher numbers indicated higher support.

Students' Assessment of Faculty Individual Accommodations

To assess instructor accommodations participants were asked "to what degree would you agree that the instructor made accommodations for differences in people's lives individually (when students asked) due to SARS-CoV-2?". Questions were rated on a 5-point Likert scale ranging from strongly agree (1) to strongly disagree (5). Answers were reversed scored so that higher numbers indicated greater accommodations.

Students' -Report of Current Grades & Prospects for Pandemic-Related Grade Drops

To assess students' perceptions of current (post-'lockdown') grades and the extent of 'lockdown'-related dips in their course grade, respondents were asked to estimate the status of their course grade both at the mid-semester point (before classes went online) and post-'lockdown.' Answer options were 10 grade categories ranging from A, then A- to F, as well as P. To assess academic achievement according to grades, passing or 'P' courses were excluded, and higher academic achievement was given higher numerical weight (A=9, F=1). The differences in these two grades were compared.

Data Analytic Plan

Hypotheses were evaluated via mixed linear models (MLM) in order to account for the hierarchical nature of the data (Raudenbush & Bryk, 2002). That is, course level data (faculty gender) (Level 1) was nested within students (Level 2). This allows for analysis of the primary variables of interest (role of faculty gender on course grades

and support) while acknowledging that course data were not independent (most participants reported on multiple current courses). All models included a random intercept to control for participant effects. That is, significant intercept effects accounts for Level 2 variance in data that differs between subjects (e.g., a student might consistently grade all course instructors more or less harshly than other students). This is the primary utility of MLM models, which provide a more conservative test of data variance than a test that ignores data dependence (e.g., *t*-test). Faculty gender, the primary predictor, was entered as a fixed effect in all models. Four MLM were run to examine whether faculty gender influenced student's perception of: faculty support, individual accommodations, mid-semester course grade (pre-pandemic) and final course grade. The model for instructor support is presented below:

Level 1 Equation

$$Instructor\ Support_i = \pi_{0i} + \pi_{1i}\ (Instructor\ Gender) + e_i$$

Level 2 Equation

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10}$$

Mixed Model

$$Instructor\ Support_i = \beta_{00} + \beta_{10}(Instructor\ Gender) + r_{0i} + e_i$$

To assess changes in perceived grades, a fifth model was run with final course grade as the dependent variable, covarying for mid-semester grade as a fixed effect on Level 1. Thus, if faculty gender had a significant fixed effect in this model, faculty gender is predicting differences in course grade from pre-pandemic to the end of the semester.

Mixed Model: *Final course grade*$_i = \beta_{00} + \beta_{10}$(Instructor Gender) $+ \beta_{20}$(Mid-semester Grade) $+ r_{0i} + e_i$.

Results

Eighty (80) undergraduate students reported on courses they were currently enrolled in, which ranged from 1—6 classes given full and part-time status. Overall, data on 362 courses was collected across participants($M = 4.53$, $SD = 1.27$; *median* $= 5$). Students reported having a (cis-)male professor for 207 courses, a (cis-)female professor for 149 courses, and a professor who was non-binary or transgender for 4 courses (faculty gender was not reported for 2 courses). Our dataset included low numbers for nonbinary or transgender instructors *and* instructors of color. Across the institution, less than 5% of the faculty body identify with a racial/ethnic identity

aside from "white." Therefore, we narrow the scope of our analytic claims exclusively to (binary, cis-) gender given the lack of variance in perceived instructor race and ethnicity, and bookmark more precise questions regarding the impact of non-binary gender, race, and other salient components of instructor identity or marginalization (arguably requiring very intentional, targeted, and well-crafted oversampling strategies) for further study.

First, we examined students' perceptions of the degree to which they felt supported by their instructor of record. The five MLM model outcomes are displayed in Table 2. Of note, all random intercepts were significant, indicating participants varied in how they tended to rate faculty (and models that had not controlled for subject effects would have had erroneously inflated effect-sizes). As anticipated, students rated their women faculty instructors as significantly more supportive than their men faculty instructors, $F(1, 333) = 7.23$, $\beta = 0.33$, $p = .008$, ($M_{\text{Women}} = 3.93$ vs. $M_{\text{Men}} = 3.60$). Then, we tested student recollections of their instructors making pandemic-related course accommodations for individual students' unique circumstances (upon request). The effect of faculty gender on individual accommodations trended to the threshold of statistical significance, $F(1, 327) = 3.09$, $\beta = 0.18$, $p = .080$, ($M_{\text{Women}} = 4.09$ vs. $M_{\text{Men}} = 3.92$).

Next, we examined whether instructor gender was associated with students' perceived academic performance (or student-reported grades). First, we examined whether their perceived mid-semester (pre-pandemic) grades differed according to faculty gender. Students did not report a difference in perceived mid-semester grades according to their instructor's gender. $F(1, 319) = 0.96$, $\beta = 0.16$, $p = .328$, ($M_{\text{Women}} = 7.77$ vs. $M_{\text{Men}} = 7.62$), with 7 and 8 corresponding with B and A- respectively. Next, we examined whether students reported significantly lower perceived current (estimated final) grades in courses with male instructors (M = 7.16, SE = .15), compared to female instructors (M = 7.57; SE = .16), $F(1, 323) = 4.92$, $\beta = 0.39$, $p = .027$. Contrastingly, instructor gender demonstrated a significant effect on estimated final grades, as students reported lower estimated final grades for courses taught by men instructors. The third model tested whether the change in perceived grades across the semester differed according to instructor gender (by covarying for mid-semester perceived grades). The effect of instructor gender was significant, $F(1, 319) = 4.69$, $\beta = 0.26$, $p = .031$, confirming that classes with male instructors reported significant drops in perceived grades across the semester whereas classes with female instructors did not.

Finally, we considered whether students' gender moderated effects by covarying for gender on the intercept and the faculty gender slope. Student's gender did not alter the effect of faculty gender in any model (lowest $p = .376$).

Discussion

In sum, our data largely aligned with our expectations: Students reported their women instructors, in contrast with their men instructors, were more supportive, and made greater individualized accommodations due to the impacts of SARS-CoV-2 (when requested). Further, students anticipated a less punitive approach to grading

Table 2 Faculty gender predicts student's perception of course support and learning outcomes

Mixed Linear Model	Support		Accommodations		Mid-semester grade		Final grade		Course grade Δ	
	β	SE	β	SE	β	SE	β	SE	β	SE
Fixed Parameters										
Level 1										
β_{00} Intercept	3.93***	0.11	4.09***	0.10	7.77***	0.16	7.54***	0.17	1.33***	0.34
β_{10} Faculty Gender	0.33**	0.12	0.18†	0.10	0.16	0.16	0.39*	0.18	0.26*	0.12
β_{20} Mid-semester grade	–	–	–	–	–	–	–	–	0.80***	0.04
Random Paramete										
r_{0i} Intercept	0.28**	0.09	0.27***	0.07	0.81***	0.20	0.87***	0.23	0.47***	0.12
e_i Residuals	1.13***	0.10	0.75***	0.06	1.88***	0.16	2.33***	0.20	1.09***	0.09

Faculty Gender: men = 0, women = 1; Course grade Δ: examining perceived change in course grade at the end of the semester by covarying for mid-semester grade. Intercept: Level 2 variance (difference in DVs according to student); Residuals: remaining Level 1 variance

*** $p < .001$, ** $p < .01$, * $p \le .05$, † $p = .080$

from their women instructors—both in their anticipated final grade outcomes and in any estimated grade performance gaps from mid-semester to final grade assignment. Put another way, students with women instructors felt more confident about the assessment of the quality of their coursework during the initial transition to remote learning. Students reported smaller anticipated (pre-pandemic to post-evacuation) grade drops in courses with women instructors than in courses taught by men, as measured by the distance between self-reported grade estimates at midterm and at the pre-final, post-evacuation moment of data collection. Notably, such effects for instructor gender were not present in students' estimates of mid-semester (pre-pandemic) grades.

Overall, the literature on gender schemas (Bachen et al., 1999) and implicit bias makes clear that gender disparities are repeatedly and reliably reproduced in workplace evaluations due to the precognitive and deeply embedded associations of femininity with nurturance and irrationality (Latu et al., 2011), which then translates to extra (uncompensated and poorly rewarded) student demands on women faculty (El-Alayli et al., 2018), as well as occasional gendered charges of incompetence and lack of objectivity in evaluating student work (Sinclair & Kunda, 2000).

We interpret our finding of higher student ratings (regarding support; individualized, pandemic-related accommodations; and sensitivity to undue grading penalties) for women faculty as likely pointing to the ongoing gendered burdens of the academic workplace. Women faculty are all-too-aware of the career penalties that can accompany a dip in any monitored performance metrics—and such a dip may be risked in any refusal to perform all of the labors required in demonstrating to students one's extraordinary levels of support, flexibility, attention, and carefully-calibrated course expectations. Further, the current data were collected during a theoretically and empirically valuable moment in time, which was marked by wholly unprecedented institutional crisis. Thus, we are afforded insights (via students' reports of their learning experiences with their instructors) into the circumstances of emergency, distress, and upheaval that may have informed women instructors' pedagogical strategies, given either their tendencies towards demonstrating greater 'academic altruism,' or greater labor performed in anticipation of the gendered punishments of bias in student evaluations. In any case, our findings should be interpreted against the backdrop of the pandemic-driven, gendered accelerations in workplace and domestic demands, declines in women faculty's research productivity, and women's greater vulnerability to mental health challenges. Though we cannot be certain that our data point to actual gendered shifts in faculty behavior, both prior research and pandemic-era studies point to an interpretation of the data that invokes gender disparate labors, crafted to respond to gender disparate workplace expectations and evaluative standards.

Certainly, more detailed, systematic, targeted research is needed to flesh the mechanisms of causation behind these phenomena out fully. Future studies could target either observed or self-reported faculty behaviors more directly. Despite any interpretive ambiguities, it is unsurprising to find that during a time of global crisis, students' reported experiences with the nurturing/emotional labors of women faculty would run steep (and perhaps for women faculty, to an entirely uncertain

degree). Given the emphasis on SETs in university performance metrics, we argue women faculty likely felt compelled to 'rise to the occasion' and give of themselves, even at the expense of their mental health (Docka-Filipek & Stone, 2021), research productivity (Cui et al., 2022; Squazzoni et al., 2020), and likely, steeper obligations to other dependents. Our findings add credibility to the claims that women faculty work harder for the same career results (O'Meara et al., 2017). We endeavor to explain these seeming empirical anomalies to better illuminate the contours of structural gender inequality (and potentially, other significant forms of power imbalance) that pervade the academy at all levels of evaluation, promotion, pay, and career security.

Further, women's higher scores on our study-specific measures should be interpreted in light of all emerging data that points to women's greater conscripted responsibility for and resulting investment in *both* the public *and* private labors of caregiving/ "service," particularly amid the pandemic. Nevertheless, any interpretation of women's scores on our measures amounts to reducing a complex unfolding process to a 'slice of time,' snapshot moment, where likely gender disparities in labor invested to produce equitable performance outcomes (measured either by SETs or students' self-reported achievements) may disappear or 'melt' into the data outcome at the endpoint of the term—resulting in a potential underestimation of gender inequities in our data, and in other similar studies (Laube et al., 2007). Therefore, time-use studies will be especially important for progress on questions regarding either career inequalities, the "leaky pipeline," or even exit from academe. Regardless, any interpretation of our findings seeking to dismiss the degree of magnitude of 'hidden service' labors would sit in contradiction with multiple decades' worth of repeatedly replicated findings on the workings of gender disparities in leadership expectations including: (Eagly & Karau, 2002), implicit bias (Cheryan & Markus, 2020; Devine, 1989), role congruity/role conflict (Eagly & Karau, 2002), cognitive/cultural schemas [or "lenses," as in Bem, 1993], the expectations and mechanisms that compel gender performativity (Butler, 1990), and the processes through which gender inequality is reproduced in complex organizations and institutions (Acker, 1990).

Data limitations include the reality that student responses were primarily drawn from the rolls in social science classes, and they therefore are an unrepresentative sample of the overall 'undergraduate population' at US institutions of higher education. It's hard to say how students' interests/elected majors may have impacted our data (signaling another avenue for future research), though it stands to reason that students enrolled in social science courses may be less inclined to report in a fashion that underscores gendered differences between their instructors, due to some degree of student awareness of (and felt moral repugnance for) the workings of gender inequity. However, this potential effect may have been countered by the fact that the student body at the university where the data was gathered tilts slightly more conservative than average, which may have compelled heightened expectations for faculty to engage especially normatively adherent gender performances. Certainly, it is plausible that conservatism in the student body may have also contributed to the finding that *students'* gender generated no significant impact on the findings, pointing to another avenue for more targeted investigation.

Further, our data was gathered at a singular institution in the Western United States, which further compromises the extent to which results may or may not generalize to other college students throughout the country. The pool of students and faculty were disproportionately cis-gendered and white (though it bears mention that on the whole, this problem of gender identity and racial homogeneity and exclusion is pervasive in U.S. higher education). Because both our student respondents and the faculty students reported on were largely homogenous (though again, not unrepresentative of the US professoriate), our sample limits our capacity to make any meaningful analytic claims that generalize beyond white, cis-gender instructors,though the literature on implicit and explicit bias also suggests that gender is interpreted in the context of one's race, further complicating or compounding student bias (Gutierrez y Muhs et al., 2012). Ultimately, the literature on role congruence dictates that women may be punished by students for deviating from gendered (and racialized—again, this is an avenue for future fruitful research) behavioral expectations, which include demands for accommodations, special attention, encouragement, etc. (Sprague & Massoni, 2005).

Lastly, we assessed students' *perceptions* of their grades, not *actual* grades. Although asking students about their anticipated final grade outcome after the issuance of final grades would have been ideal, widespread data collection would have likely proven near impossible, not only because the academic year had concluded, but also due to the constraints and pressures imposed by the pandemic. Nevertheless, we argue that in some ways, student perceptions are more valuable than final grade realities, as the former are empirically proven to be more influential in driving SET outcomes than actual learning or student achievement (Boring & Ottoboni, 2016; Uttl et al., 2017).

Despite the aforementioned limitations, we nevertheless believe our data points conclusively to the gender asymmetries in students' expectations for their instructors' pedagogical strategies and overall investments. To the degree that such asymmetries impact SETs, and metrics derived from SETs impact performance assessments, and then by extension, career outcomes, such asymmetries may also drive widespread gender inequities across the academy. Further, at the 'bird's eye' level, we argue that instead of an 'unfortunate' accident, any gender bias in SETs is one reliable mechanism through which the neoliberal university extracts substantial amounts of unpaid emotional labor from marginalized faculty [see Padilla, 1994], conscripting them into chronologically and psychologically taxing teaching activities that come with scant tangible rewards for the concrete advancement of their careers. Ultimately, this extractive trend of securing the deepest time-intensive and emotionally invested commitments from scholars most vulnerable to the negative impacts of the pandemic constitutes a harmful pattern which may stand to accelerate as the COVID crisis wears on. The pandemic may be cumulative in its negative impacts, a reality we must consider as any 'sunset' on this extended crisis reaches farther and farther out onto the horizon. Further, the crisis of the pandemic may be used as a pretext for neoliberal university restructuring, in line with one Colorado administrator's admonition to "never waste a good pandemic" (Flaherty, 2020).

Fruitful avenues for future research, which may permit stronger or sharper claims making, includes more close examination of students' gendered expectations for

faculty's pedagogical investments assessments alongside detailed time-use diaries, observations of faculty behavior, or other self-reporting. Such studies may help to determine how faculty's teaching labors vary alongside their other time investments in research and in varying forms of service. Additionally, further studies can and should be conducted with larger and more diverse samples in order to examine any potential moderating effects (such as teaching load, academic profile of the student body, gender/race composition of the faculty, political/attitudinal/demographic leanings of the student body, etc.).

Conclusion

The pattern of results support that women faculty were evaluated by their students as more supportive, accommodating, flexible, lenient, and less punitive in a time of great crisis serves as a testament to the survival strategies women faculty (and potentially, other underrepresented scholars) have had to engage in order to navigate the empirically confirmed, deeply embedded, and now potentially amplifying possibility for inequality in career outcomes on the basis of gender. Further, such findings should be interpreted against the backdrop of all the *additional* gendered/racialized pressures the pandemic has amplified. Together, these realities have assembled a recipe for a potentially widespread pushout of underrepresented scholars (Malisch et al., 2020).

For these reasons, feminist and race critical scholars must push for a number of institutional reforms in the ways faculty teaching and service labors are evaluated. First, "secret service" must be drug into the light, measures developed, and rewards codified (see suggested proposals in Domingo et al., 2022), as such labors are likely the glue that holds institutions together during precarious times—especially if such precarity is fueled, at least in part, by declining enrollments due to student attrition. Conceivably, "unsupported" students would be at greater risk for academic exit. Additionally, institutions should consider the reinterpretation or even outright elimination of the use of SETs in promotion, tenure, and performance evaluations—indeed, a number of institutions (most notably in California and Oregon) adopted this strategy in years leading up to the pandemic, amid mounting institutional concerns about liability for practices empirically proven to play a direct role in causing discriminatory, disparate impact (Flaherty, 2018). Other institutions have chosen to 'downgrade' or reinterpret the meaning of SETs amid the pressures and disparate impacts of the COVID-19 pandemic (Lederman, 2020; Mickey et al., 2022). Some schools may consider a modification of the framing of their institutional proctoring/delivery of the task of SETs to students in order to reduce the incidence of gender bias. Recently proposed, evidence-based strategies include anti-bias training for students prior to SET administration (Peterson et al., 2019) (though overall the literature demonstrates mixed evidence on efficacy and possible 'backlash' effects, suggesting caution); a reduction of scalar evaluative options (Rivera & Tilcsik, 2019) on such instruments; implementing self-affirming reflective exercises for students to complete prior to proctoring SETs [which seemingly 'deflate' the scores of men instructors to relative

parity with women faculties' scores (Hoorens et al., 2021)]; or the renaming of such questionnaires to "student experience reports" or "student learning impressions", as some analysts argue that students are under qualified to evaluate effective pedagogy (see Kreitzer & Sweet-Cushman, 2021). Further, incentives for student participation that are meted out at the institutional level to boost completion rates may serve to blunt some of the most extreme variation in SET responses for women faculty, thereby moderating any disproportionately negative scores. University authorities would also do well to consider a push for the consideration of multiple artifacts of teaching, selected to both point to effective instruction and/or the need for pedagogical improvement. Holistic and well-rewarded peer evaluations performed only by colleagues with some training in inter/transdisciplinary best practices from the scholarship of teaching and learning may prove useful. Anti-bias trainings for those university authorities tasked with translating SETs into faculty career judgements may compel more nuanced interpretation.

Ultimately, one of the most important lessons of the pandemic has been its capacity to drag 'hidden' labors once less-recognized, or even invisible, into the light of day. Therefore, a shifting of faculty performance metrics to both minimize the impact of students bias on the career trajectories of marginalized faculty *and* reward 'hidden service,' may represent a step towards rewarding precisely the type of redistributive, prosocial, and collectivist behaviors the US academy will need to weather the storm of the current crisis. Ultimately, higher education in the US may well require institutional reevaluations of the interpretation and use of performance metrics that have been repeatedly, empirically demonstrated to punish and push out the cadre of faculty who do the lion's share of the necessary, hard work of rendering the experience of higher education (and relatedly, the mission of liberal arts inquiry and a corresponding critical interrogation of the self and the social order) less a coldly 'efficient' transactional experience, and more a transformational, deeply rewarding journey of exploration of the self and one's place in the broader moral order of the world. Put another way: If faculty are not adequately rewarded, or perhaps indeed, continue to be inadvertently punished for taking the time to approach their students flexibly, supportively, and sensitively, not only will the demographic composition of the professoriate become increasingly homogenous and unidimensional, so too will our standards for 'excellence,' growth, 'rigor,' and care. Academe undervalues 'service' at its own peril. Performance metrics in higher education must therefore adapt as the world changes, not only to navigate the crises presented by the pandemic, but to emerge from them as spaces where new, as-yet undetermined forms of mutualism, care, and reciprocity remain possible.

Data Availability Data is available upon request from the author/s.

Declarations

Ethics Approval The questionnaire and methodology for this study was approved by the Institutional Review Board of Utah State University (approvals notification/documentation available upon request). Informed consent was secured from all survey participants.

Competing Interests The authors have no competing interests to declare.

References

Acker, J. (1990). Hierarchies, jobs, bodies: A theory of gendered organizations. *Gender & Society*, *4*(2), 139–158. https://doi.org/10.1177/089124390004002002

Acker, S., & Feuerverger, G. (1996). Doing good and feeling bad: The work of women university teachers. *Cambridge Journal of Education*, *26*(3), 401–422. https://doi.org/10.1080/0305764960 260309

Arbuckle, J., & Williams, B. D. (2003). Students' perceptions of expressiveness: Age and gender effects on teacher evaluations. *Sex Roles*, *49*(9), 507–516. https://doi.org/10.1023/A:1025832707002

Ashencaen Crabtree, S., & Shiel, C. (2019). "Playing mother": Channeled careers and the construction of gender in academia. *SAGE Open*, *9*(3), 2158244019876285. https://doi.org/10.1177/2158244019876285

Babcock, L., Recalde, M. P., Vesterlund, L., & Weingart, L. (2017). Gender differences in accepting and receiving requests for tasks with low promotability. *American Economic Review*, *107*(3), 714–747. https://doi.org/10.1257/aer.20141734

Bachen, C. M., McLoughlin, M. M., & Garcia, S. S. (1999). Assessing the role of gender in college students' evaluations of faculty. *Communication Education*, *48*(3), 193–210. https://doi.org/10.1080/03634529909379169

Bagilhole, B., & Goode, J. (1998). The 'Gender Dimensions' of both the 'Narrow' and 'Broad' curriculum in UK higher education: Do women lose out in both? *Gender and Education*, *10*(4), 445–458. https://doi.org/10.1080/09540259820862

Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology*, *79*(3), 308. https://doi.org/10.1037/0022-0663.79.3.308

Bem, S. L. (1993). *The lenses of gender: Transforming the debate on sexual inequality*. Yale University Press.

Benozzo, A., Koro, M., Vasquez, A., Vitrukh, M., Barbetta, P., & Long, C. (2022). A femin… manifesto: Academic ecologies of care and cure during a global health pandemic. *Gender, Work and Organization*, *29*(4), 1236–1258. https://doi.org/10.1111/gwao.12762

Boncori, I. (2020). The never-ending shift: A feminist reflection on living and organizing academic lives during the coronavirus pandemic. *Gender, Work and Organization*, *27*(5), 677–682. https://doi.org/10.1111/gwao.12451

Boring, A., & Ottoboni, K. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *Science Open Research*. https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1

Branch, E.H. (Ed.). (2017). *Pathways, potholes, and the persistence of women in science: Reconsidering the pipeline*. Lexington Books.

Branicki, L. J. (2020). COVID-19, ethics of care and feminist crisis management. *Gender, Work and Organization*, *27*(5), 872–883. https://doi.org/10.1111/gwao.12491.

Brown, W. (2015). *Undoing the demos: Neoliberalism's stealth revolution*. MIT Press.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.

Calarco, J. M., Meanwell, E., Anderson, E., & Knopf, A. (2020). "My husband thinks I'm crazy": COVID-19-related conflict in couples with young children. https://doi.org/10.31235/osf.io/cpkj6

Calarco, J. M., Meanwell, E., Anderson, E. M., & Knopf, A. S. (2021). By default: How mothers in different-sex dual-earner couples account for inequalities in pandemic parenting. *Socius*, 7, 23780231211038783. https://doi.org/10.31235/osf.io/hgnfs

Cate, L., Ward, L. W. M. & Ford, K. S. (2022). Strategic ambiguity: How pre-tenure faculty negotiate the hidden rules of academia. *Innovative Higher Education, 47*, 795–812. https://doi.org/10.1007/s10755-022-09604-x

Centra, J. A. (2000). Evaluating the teaching portfolio: A role for colleagues. *New Directions for Teaching and Learning*, *83*, 87–93. https://doi.org/10.1002/tl.8307

Cheryan, S., & Markus, H. R. (2020). Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychological Review*, *127*(6), 1022. https://doi.org/10.1037/rev0000209

CohenMiller, A., Izekenova, Z. (2022). Motherhood in academia during the COVID-19 pandemic: An international online Photovoice study addressing issues of equity and inclusion in higher education. *Innovative Higher Education, 47*, 813–835. https://doi.org/10.1007/s10755-022-09605-w

Collins, C., Landivar, L. C., Ruppanner, L., & Scarborough, W. J. (2021). COVID-19 and the gender gap in work hours. *Gender, Work and Organization*, *28*, 101–112. https://doi.org/10.1111/gwao.12506

Chism, N., & Chism, G. W. (2007). *Peer review of teaching: A sourcebook*. Jossey Bass.

Craig, L., & Churchill, B. (2021). Dual-earner parent couples' work and care during COVID-19. *Gender, Work and Organization*, *28*, 66–79. https://doi.org/10.1111/gwao.12497

Cui, R., Ding, H., & Zhu, F. (2022). Gender inequality in research productivity during the COVID-19 pandemic. *Manufacturing & Service Operations Management*, *24*(2), 707–726. https://doi.org/10.1287/msom.2021.0991

Davis, J. C., Li, E. P. H., Butterfield, M. S., DiLabio, G. A., Santhagunam, N., & Marcolin, B. (2022). Are we failing female and racialized academics? A Canadian national survey examining the impacts of the COVID-19 pandemic on tenure and tenure-track faculty. *Gender, Work and Organization*, *29*(3), 703–722. https://doi.org/10.1111/gwao.12811

Del Boca, D., Oggero, N., Profeta, P., & Rossi, M. (2020). Women's and men's work, housework and childcare, before and during COVID-19. *Review of Economics of the Household*, *18*(4), 1001–1017. https://doi.org/10.1007/s11150-020-09502-1

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5. https://doi.org/10.1037/0022-3514.56.1.5

DiQuinzio, P. (1993). Exclusion and essentialism in feminist theory: The problem of mothering. *Hypatia*, *8*(3), 1–20. https://doi.org/10.1111/j.1527-2001.1993.tb00033.x

Docka-Filipek, D., & Stone, L. B. (2021). Twice a "housewife": On academic precarity, "hysterical" women, faculty mental health, and service as gendered care work for the "university family" in pandemic times. *Gender, Work and Organization*, *28*(6), 2158–2179. https://doi.org/10.1111/gwao.12723

Domingo, C. R., Gerber, N. C., Harris, D., Mamo, L., Pasion, S. G., Rebanal, R. D., & Rosser, S. V. (2022). More service or more advancement: Institutional barriers to academic success for women and women of color faculty at a large public comprehensive minority-serving state university. *Journal of Diversity in Higher Education*, *15*(3), 365. https://doi.org/10.1037/dhe0000292

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573. https://doi.org/10.1037/0033-295X.109.3.573

Eddleston, K. A., & Mulki, J. (2017). Toward understanding remote workers' management of work–family boundaries: The complexity of workplace embeddedness. *Group & Organization Management*, *42*(3), 346–387. https://doi.org/10.1177/1059601115619548

El-Alayli, A., Hansen-Brown, A. A., & Ceynar, M. (2018). Dancing backwards in high heels: Female professors experience more work demands and special favor requests, particularly from academically entitled students. *Sex Roles*, *79*(3), 136–150. https://doi.org/10.1007/s11199-017-0872-6

Esarey, J., & Valdes, N. (2020). Unbiased, reliable, and valid student evaluations can still be unfair. *Assessment & Evaluation in Higher Education*, *45*(8), 1106–1120. https://doi.org/10.1080/02602938.2020.1724875

Ewing-Nelson, C. (2021). All of the jobs lost in december were women's jobs. *National Women's Law Center, 3*. https://nwlc.org/wp-content/uploads/2021/01/December-Jobs-Day.pdf

Fan, Y., Shepherd, L. J., Slavich, E., Waters, D., Stone, M., Abel, R., & Johnston, E. L. (2019). Gender and cultural bias in student evaluations: Why representation matters. *PLoS One*, *14*(2), e0209749. https://doi.org/10.1371/journal.pone.0209749

Fingerhut, H. (2017). Republicans skeptical of colleges' impact on US, but most see benefits for workforce preparation. https://www.pewresearch.org/fact-tank/2017/07/20/republicans-skeptical-of-colleges-impact-on-u-s-but-most-see-benefits-for-workforce-preparation/

Flaherty, C. (2018). Teaching eval shake-up. *Inside Higher Ed*, *22*. https://www.insidehighered.com/news/2018/05/22/most-institutions-say-they-value-teaching-how-they-assess-it-tells-different-story

Flaherty, Colleen. (2020). 'Never waste a good pandemic.' *Inside Higher Ed*. https://www.insidehighered.com/news/2020/12/04/boulder-arts-and-sciences-dean-wants-build-back-faculty-post-pandemic-one-non-tenure

Freeman, H. R. (1994). Student evaluations of college instructors: Effects of type of course taught, instructor gender and gender role, and student gender. *Journal of Educational Psychology*, *86*(4), 627. https://doi.org/10.1037/0022-0663.86.4.627

Geiger, R. (2010). Impact of the financial crisis on higher education in the United States. *International Higher Education*, (*59*). https://doi.org/10.6017/ihe.2010.59.8486

Giroux, H. A. (2014). Austerity and the poison of neoliberal miseducation. *symplokē*, *22*(1–2), 9–21.

Górska, A. M., Kulicka, K., Staniszewska, Z., & Dobija, D. (2021). Deepening inequalities: What did COVID-19 reveal about the gendered nature of academic work?. *Gender, Work and Organization*, *28*(4), 1546–1561. https://doi.org/10.1111/gwao.12696

Gregg, M. (2011). *Work's intimacy*. Polity Press.

Grosz, E. (1995). *Space, time, and perversion*. Routledge.

Guarino, C. M., & Borden, V. M. (2017). Faculty service loads and gender: Are women taking care of the academic family? *Research in Higher Education*, *58*(6), 672–694. https://doi.org/10.1007/s11162-017-9454-2

Muhs, G. Gy., Niemann, Y. F., González, C. G., Harris, A. P. (Eds.). (2012). *Presumed incompetent: The intersections of race and class for women in academia*. University Press of Colorado.

Hanasono, L. K., Broido, E. M., Yacobucci, M. M., Root, K. V., Peña, S., & O'Neil, D. A. (2019). Secret service: Revealing gender biases in the visibility and value of faculty service. *Journal of Diversity in Higher Education*, *12*(1), 85. https://doi.org/10.1037/dhe0000081

Heffernan, T. (2022). Sexism, racism, prejudice, and bias: A literature review and synthesis of research surrounding student evaluations of courses and teaching. *Assessment & Evaluation in Higher Education*, *47*(1), 144–154. https://doi.org/10.1080/02602938.2021.1888075

Hoorens, V., Dekkers, G., & Deschrijver, E. (2021). Gender bias in student evaluations of teaching: Students' self-affirmation reduces the bias by lowering evaluations of male professors. *Sex Roles*, *84*(1), 34–48. https://doi.org/10.1007/s11199-020-01148-8

Huang, J., Gates, A. J., Sinatra, R., & Barabási, A. L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, *117*(9), 4609–4616. https://doi.org/10.1073/pnas.1914221117

Huston, T. A. (2006). Race and gender bias in higher education: Could faculty course evaluations impede further progress toward parity? *Seattle Journal for Social Justice*, *4*(2), 34.

Johnson, S. (2022). Women deserve better: A discussion on COVID-19 and the gendered organization in the new economy. *Gender, Work and Organization*, *29*(2), 639–649. https://doi.org/10.1111/gwao.12785

Kasymova, S., Place, J. M. S., Billings, D. L., & Aldape, J. D. (2021). Impacts of the COVID-19 pandemic on the productivity of academics who mother. *Gender, Work and Organization*, *28*, 419–433. https://doi.org/10.1111/gwao.12699

Kim, D. Y. (2003). Voluntary controllability of the implicit association test (IAT). *Social Psychology Quarterly*, *83–96*. https://doi.org/10.2307/3090143

Kogan, V., Genetin, B., Chen, J., & Kalish, A. (2022). Students' grade satisfaction influences evaluations of teaching: Evidence from individual-level data and an experimental intervention. https://www.edworkingpapers.com/ai22-513

Key, E., & Ardoin, P. (2019). *Students rate male instructors more highly than female instructors. We tried to counter that hidden bias*. Washington Post. https://www.washingtonpost.com/politics/2019/08/20/students-rate-male-instructors-more-highly-than-female-instructors-we-tried-to-counter-that-hidden-bias/

Kreitzer, R. J., & Sweet-Cushman, J. (2021). Evaluating student evaluations of teaching: A review of measurement and equity bias in SETs and recommendations for ethical reform. *Journal of Academic Ethics*, *1–12*. https://doi.org/10.1007/s10805-021-09400-w

Latu, I. M., Stewart, T. L., Myers, A. C., Lisco, C. G., Estes, S. B., & Donahue, D. K. (2011). What we "say" and what we "think" about female managers: Explicit versus implicit associations of women with success. *Psychology of Women Quarterly*, *35*(2), 252–266. https://doi.org/10.1177/0361684310383811

Laube, H., Massoni, K., Sprague, J., & Ferber, A. L. (2007). The impact of gender on the evaluation of teaching: What we know and what we can do. *NWSA Journal*, 87–104.

Lederman, D. (2020). Evaluating teaching during the pandemic. *Inside Higher Ed*, *8*(April). https://www.insidehighered.com/digital-learning/article/2020/04/08/many-colleges-are-abandoning-or-downgrading-student-evaluations

Malisch, J. L., Harris, B. N., Sherrer, S. M., Lewis, K. A., Shepherd, S. L., McCarthy, P. C., ... & Deitloff J. (2020). In the wake of COVID-19, academia needs new solutions to ensure gender equity. *Proceedings of the National Academy of Sciences*, *117*(27), 15378–15381. https://doi.org/10.1073/pnas.2010636117

Mengel, F., Sauermann, J., & Zölitz, U. (2019). Gender bias in teaching evaluations. *Journal of the European Economic Association*, *17*(2), 535–566. https://doi.org/10.1093/jeea/jvx057

Mickey, E. L., Misra, J., & Clark, D. (2022). The persistence of neoliberal logics in faculty evaluations amidst Covid-19: Recalibrating toward equity. *Gender, Work and Organization* https://doi.org/10.1111/gwao.12817

Miller, C. C. (2020). *Nearly half of men say they do most of the home schooling. 3 percent of women agree*. The New York Times, 6. https://www.nytimes.com/2020/05/06/upshot/pandemic-chores-homeschooling-gender.html

Miller, J. E., & Seldin, P. (2014). Changing practices in faculty evaluation. *Academe*, *100*(3), 35–38.

Misra, J., Lundquist, J. H., & Templer, A. (2012). *Gender, work time, and care responsibilities among faculty 1. In Sociological Forum* (Vol. 27, 2, pp. 300–323). Blackwell Publishing Ltd. https://doi.org/10.1111/j.1573-7861.2012.01319.x

Murray, H. G. (1984). The impact of formative and summative evaluation of teaching in north American universities. *Assessment and Evaluation in Higher Education*, *9*(2), 117–132. https://doi.org/10.1080/0260293840090204

Newport, F., & Busteed, B. (2017). *Why are republicans down on higher Ed?*. Gallup News. https://news.gallup.com/poll/216278/why-republicans-down-higher.aspx

Özkazanç-Pan, B., & Pullen, A. (2020). Gendered labour and work, even in pandemic times. *Gender, Work and Organization*, *27*(5), 675. https://doi.org/10.1111/gwao.12516

O'Meara, K., Kuvaeva, A., Nyunt, G., Waugaman, C., & Jackson, R. (2017). Asked more often: Gender differences in faculty workload in research universities and the work interactions that shape them. *American Educational Research Journal*, *54*(6), 1154–1186. https://doi.org/10.3102/0002831217716767

Padilla, A. M. (1994). Research news and comment: Ethnic minority scholars; research, and mentoring: Current and future issues. *Educational Researcher*, *23*(4), 24–27. https://doi.org/10.3102/0013189X023004024

Pereira, M. D. M. (2021). Researching gender inequalities in academic labor during the COVID-19 pandemic: Avoiding common problems and asking different questions. *Gender, Work and Organization*, *28*, 498–509. https://doi.org/10.1111/gwao.12618

Peterson, D. A., Biederman, L. A., Andersen, D., Ditonto, T. M., & Roe, K. (2019). Mitigating gender bias in student evaluations of teaching. *PLoS One*, *14*(5), e0216241. https://doi.org/10.1371/journal.pone.0216241

Petts, R. J., Carlson, D. L., & Pepin, J. R. (2021). A gendered pandemic: Childcare, homeschooling, and parents' employment during COVID-19. *Gender, Work and Organization*, *28*, 515–534. https://doi.org/10.1111/gwao.12614

Piatak, J., & Mohr, Z. (2019). More gender bias in academia? Examining the influence of gender and formalization on student worker rule following. *Journal of behavioral Public Administration*, *2*(2). https://doi.org/10.30636/jbpa.22.76

Pittman, C. T. (2010). Race and gender oppression in the classroom: The experiences of women faculty of color with white male students. *Teaching Sociology*, *38*(3), 183–196. https://doi.org/10.1177/0092055X10370120

Plotnikof, M., & Utoft, E. H. (2022). The "new normal" of academia in pandemic times: Resisting toxicity through care. *Gender, Work and Organization*, *29*(4), 1259–1271. https://doi.org/10.1111/gwao.12778

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). SAGE. http://us.sagepub.com/en-us/nam/hierarchical-linear-models/book9230

Régner, I., Thinus-Blanc, C., Netter, A., Schmader, T., & Huguet, P. (2019). Committees with implicit biases promote fewer women when they do not believe gender bias exists. *Nature Human Behaviour*, *3*(11), 1171–1179. https://doi.org/10.1038/s41562-019-0686-3

Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors. Com. *Journal of Diversity in Higher Education*, *3*(3), 137. https://doi.org/10.1037/a0019865

Renström, E. A., Gustafsson Sendén, M., & Lindqvist, A. (2021). *Gender stereotypes in student evaluations of teaching. In Frontiers in education* (p. 280). Frontiers. https://www.frontiersin.org/article/10.3389/feduc.2020.571287

Richards, B. N. (2019). *9 Faculty assessments as tools of oppression. In Intersectionality and higher education* (pp. 136–151). Rutgers University Press.

Rivera, L. A., & Tilcsik, A. (2019). Scaling down inequality: Rating scales, gender bias, and the architecture of evaluation. *American Sociological Review*, *84*(2), 248–274. https://doi.org/10.1177/0003122419833601

Seldin, P., Miller, J. E., & Seldin, C. A. (2010). *The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions*. John Wiley & Sons.

Shreffler, M. B., Shreffler, J. R., & Murfree, J. R. (2019). Barriers to advancement of female faculty in higher education: An examination of student evaluations of teaching, work-family conflict and perceived organizational support. *Journal of Higher Education Theory and Practice*, *19*(8), 72–85. https://doi.org/10.33423/jhetp.v19i8.2673

Sinclair, L., & Kunda, Z. (2000). Motivated stereotyping of women: She's fine if she praised me but incompetent if she criticized me. *Personality and Social Psychology Bulletin*, *26*(11), 1329–1342. https://doi.org/10.1177/0146167200263002

Sprague, J., & Massoni, K. (2005). Student evaluations and gendered expectations: What we can't count can hurt us. *Sex Roles*, *53*(11), 779–793. https://doi.org/10.1007/s11199-005-8292-4

Squazzoni, F., Bravo, G., Grimaldo, F., Garcıa-Costa, D., Farjam, M., & Mehmani, B. (2020). Only second-class tickets for women in the COVID-19 race. A study on manuscript submissions and reviews in 2329 Elsevier journals. A study on manuscript submissions and reviews in, 2329. https://doi.org/10.2139/ssrn.3712813

Stroebe, W. (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), 276–294. https://doi.org/10.1080/01973533.2020.1756817

Thibaut, F., & van Wijngaarden-Cremers, P. J. (2020). Women's mental health in the time of Covid-19 pandemic. *Frontiers in Global Women's Health*, *1*, 588372. https://www.frontiersin.org/article/10.3389/fgwh.2020.588372

Tuck, J. (2018). "I'm nobody's mum in this university": The gendering of work around student writing in UK higher education. *Journal of English for Academic Purposes*, *32*, 32–41. https://doi.org/10.1016/j.jeap.2018.03.006

U. S. Bureau of Labor Statistics. (2021). Current employment statistics. Table B-5. employment of women on nonfarm payrolls by industry sector, seasonally adjusted, Series Code: CES0000000010. https://www.bls.gov/webapps/legacy/cesbtab5.htm

Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, *54*, 22–42. https://doi.org/10.1016/j.stueduc.2016.08.007

Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, *23*(2), 191–212. https://doi.org/10.1080/0260293980230207

Weisshaar, K. (2017). Publish and perish? An assessment of gender gaps in promotion to tenure in academia. *Social Forces*, *96*(2), 529–560. https://doi.org/10.1093/sf/sox052

West, M. S., & Curtis, J. W. (2006). *AAUP faculty gender equity indicators 2006* (p. 85). American Association of University Professors. https://www.aaup.org/reports-publications/publications/see-all/aaup-faculty-gender-equity-indicators-2006

Yavorsky, J. E., Kamp Dush, C. M., & Schoppe-Sullivan, S. J. (2015). The production of inequality: The gender division of labor across the transition to parenthood. *Journal of Marriage and Family*, *77*(3), 662–679. https://doi.org/10.1111/jomf.12189

Yildirim, T. M., & Eslen-Ziya, H. (2021). The differential impact of COVID-19 on the work conditions of women and men academics during the lockdown. *Gender, Work and Organization*, *28*, 243–249. https://doi.org/10.1111/gwao.12529

Young, S., Rush, L., & Shaw, D. (2009). Evaluating gender bias in ratings of university instructors' teaching effectiveness. *International Journal for the Scholarship of Teaching and Learning*, *3*(2), n2. https://doi.org/10.20429/ijsotl.2009.030219

Zamarro, G., Perez-Arce, F., & Prados, M. J. (2020). *Gender differences in the impact of COVID-19*. KTLA. https://ktla.com/wp-content/uploads/sites/4/2020/06/ZamarroGenderDiffImpactCOVID-19_061820-2.pdf. Accessed 16 Jul 2021

Danielle Docka-Filipek is an Assistant Professor of Sociology at the University of Minnesota, Duluth. Her teaching expertise lies in race/class/gender inequalities; feminist theory/praxis; poverty & social welfare policy; sex, gender, sexuality, & embodiment/affect; sociology of mental health/illness; globalization; family/marriage; religion & social movements; culture, law & organizations; and teaching, learning, & inequality in higher education. Her current projects include examining racial discourse & practices in U.S. higher education via faculty and student encounters with the twin concepts of 'diversity' and 'academic freedom;' and first-generation students' & faculty members' well-being/mental health, work/family conflicts, & teaching/learning burdens amid the disruptions of the COVID-19 pandemic. She especially enjoys involving students in her research projects as both assistants and partners, and her university service/activist projects have spanned topics such as sexual violence/assault & consent-based education, access/inclusion for historically underrepresented groups, and interdisciplinary program development.

Crissa Draper is an Assistant Professor at Utah State University, and the co-director of the Psychology Department Undergraduate Program. Dr. Levin has taught exclusively online undergraduate courses, including research methods and pseudoscience, since beginning her faculty appointment in 2017. Dr. Levin implements evidence-based teaching methods of in her courses, and works to teach students in meaningful ways about science. She also co-runs an active distance undergraduate research lab, where research focuses on a variety issues that may interest undergraduate students. In addition, she regularly does research on the scholarship of teaching and learning.

Janice Snow is a social work graduate student at Utah State University. She is involved in research projects on all levels, working with faculty to complete their research, mentoring other students, and working on several of her own impactful research studies, most notably on inequities amidst the COVID-19 pandemic, including student mental health and academic challenges. Further, Janice is actively involved in the USU community through her role as the president of the MSW Student Association, and her work as an intern therapist with vulnerable populations. She has experience with working with youth in a group home setting working on helping them to control their emotions, and she also volunteers as a crisis counselor through Crisis Textline, where she helps people work on coping skills.

Lindsey B. Stone is an Associate Professor of Clinical Psychology at Georgia Southern University. Dr. Stone's research targets interpersonal mechanisms of depression risk during adolescence and emerging adulthood. She applies behavioral and psychophysiological assessments to clarify how social support strategies impact emotion regulation. Since March 2020, Dr. Stone has pivoted to chronicling the disparate impacts of the COVID-19 pandemic on depression risk; detailing the structural barriers that threaten the well-being and success of minority individuals and women in the workplace.

Authors and Affiliations

Danielle Docka-Filipek[1] [ORCID] **· Crissa Draper[2] · Janice Snow[3] · Lindsey B. Stone[4]**

[1]   Department of Studies in Justice, Culture, and Social Change, University of Minnesota, Cina Hall 228, 1123 University Drive, Duluth, MN 55812, USA

[2]   Department of Psychology, Utah State University, Logan, UT, USA

[3]   Department of Social Work, Utah State University, Logan, UT, USA

[4]   Department of Psychology, Georgia Southern University, Statesboro, GA, USA