Scientific Research Publishing

Open Journal of Social Sciences, 2026, 14(11), 269-299
https://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

The Struggles within: Mapping Perceived Inequity, Unequal Workload, and Constrained Voice among Higher Education Faculty

Maria Matthews[1], Annamarie Sisson[2], Nicholous Deal[1], Rhonda Bursey[1], Sheri Landry[1]

[1]Department of Business & Tourism, Mount Saint Vincent University, Halifax, Canada
[2]Business School, Napier University, Edinburgh, UK
Email: maria.matthews2@msvu.ca

How to cite this paper: Matthews, M., Sisson, A., Deal, N., Bursey, R., & Landry, S. (2026). The Struggles within: Mapping Perceived Inequity, Unequal Workload, and Constrained Voice among Higher Education Faculty. *Open Journal of Social Sciences, 14,* 269-299.
https://doi.org/10.4236/jss.2026.1411015

Received: October 10, 2026
Accepted: November 14, 2026
Published: November 17, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Open Access

Abstract

This exploratory cross-sectional study maps faculty-reported concerns about teaching evaluation, workload, unequal treatment, and constrained voice. Of 168 survey records, three with incompatible employment-context evidence were excluded from record-level analyses, leaving 165 records of incompletely verified faculty status. In the initial wave, 53 of 67 valid respondents (79.1%) reported evaluation-related behavioural changes; a distinct counterfactual item in the second wave received four Yes answers among six respondents. Differential treatment was reported by 47 of 72 respondents (65.3%) answering that item. Across four open-ended questions, 243 candidate answers from 73 records were identified; after respondent- and answer-level exclusions, 84 answers from 32 distinct contributors were retained for manifest content analysis. The four descriptive categories addressed evaluation-related adaptation, workload and research displacement, unequal treatment and favouritism, and gatekeeping, mentoring, voice, and strain. These are reports of perceived concerns, not population prevalence or demonstrated causal processes. Conditional belonging is proposed as a provisional account of how faculty may perceive opportunities and recognition as contingent on conformity to informal organizational expectations (Acker, 2006; Morrison, 2014; Settles et al., 2024). The findings support attention to transparent workload allocation, contextualized teaching evaluation, explicit advancement criteria, mentoring, and credible reporting mechanisms.

Keywords

Higher Education, Faculty Workload, Organizational Inequality, Employee Silence

1. Introduction

Higher education frequently presents itself as a meritocratic space in which intellectual contribution, teaching effectiveness, and service determine professional advancement (Clarke et al., 2024). Formal policies, however, do not operate independently from institutional hierarchies, discretionary decisions, and unequal access to resources (Acker, 2006). Organizational inequality can become embedded in everyday practices through work allocation, recognition, evaluation, and opportunity structures (O'Meara et al., 2019; Settles et al., 2024). These dynamics are especially consequential where professional advancement depends on formal criteria interpreted within institutional structures; research on associate professors, for example, documents substantial concern about the clarity of promotion expectations (Kulp et al., 2019).

Research on tenured associate professors shows that expectations for promotion to full professor are not always clear (Kulp et al., 2019). The present study considers whether uncertainty about expectations also appears in accounts from faculty across appointment types. The relative value assigned to research productivity, teaching effectiveness, collegiality, student responsiveness, and service may not always be transparent within departmental systems (O'Meara et al., 2019). Research has also shown that service work is not distributed evenly and that women and faculty of colour may carry greater internal service, mentoring, and diversity-related responsibilities (Guarino & Borden, 2017; Guillaume & Apodaca, 2022). These responsibilities can support students and institutions while reducing time available for research and other activities that are rewarded in advancement decisions (O'Meara et al., 2017). Department-level practices such as workload transparency, clarity, and perceived fairness are associated with faculty workload satisfaction and perceptions of equity (O'Meara et al., 2019).

Faculty experiences of inequality are also epistemic (Settles et al., 2022). Academic organizations help determine which questions, methods, journals, and forms of contribution are treated as legitimate (Settles et al., 2021). Epistemic exclusion occurs when scholars or their work are devalued through formal evaluation systems and informal interactions, particularly when disciplinary assumptions combine with identity-based bias (Settles et al., 2024). When those raising concerns believe that speaking may affect their careers, self-censorship can become a rational response to perceived interpersonal or career risk (Detert & Edmondson, 2011).

The language of oppression must therefore be used carefully (Acker, 2006; Ray, 2019). An unpleasant interaction, demanding workload, or ineffective manager is not by itself evidence of an organizational inequality regime (Acker, 2006). Systemic oppression is more appropriately examined through recurring organizational mechanisms that distribute resources, credibility, voice, and career risk unequally (Ray, 2019; Settles et al., 2024). In higher education, relevant mechanisms documented in the literature include unclear advancement expectations (Kulp et al., 2019), gendered service allocation (Guarino & Borden, 2017), bias in student

evaluations (Fan et al., 2019; MacNell et al., 2015), and scholarly devaluation through formal and informal evaluation (Settles et al., 2021).

A substantial body of literature addresses faculty inequity, cultural taxation, biased evaluation, and marginalization across different dimensions of academic work (Fan et al., 2019; Guillaume & Apodaca, 2022; O'Meara et al., 2017; Settles et al., 2021, 2024). The present study brings concerns about teaching, workload, treatment, mentoring, voice, well-being, and advancement together within one exploratory survey (see Method). Its purpose was to identify and organize faculty-reported concerns while distinguishing broad patterns of mention from the small number of accounts that provided contextual depth (see Method). The guiding research question was: What organizational concerns do higher education faculty report in relation to evaluation, workload, treatment, voice, well-being, and career progression? (see Method).

2. Literature Review

2.1. Systems of Oppression, Organizational Power, and Manipulation

Oppression is not limited to explicitly discriminatory actions (Acker, 2006). Acker's (2006) concept of inequality regimes explains how interconnected organizational practices maintain inequalities of gender, race, and class even when institutional policies appear neutral. Ray (2019) similarly argues that organizations are not race-neutral structures because rules, hierarchies, and resources can reproduce racialized outcomes. Together, these perspectives shift attention from individual intent toward the organizational processes through which apparently ordinary procedures can produce unequal outcomes (Clarke et al., 2024; Ray, 2019).

Leadership behaviour can be examined alongside organizational structures, while recognizing the settings in which the available evidence was generated. In a survey of employees in Pakistani small and medium enterprises, despotic leadership and workplace incivility were negatively associated with psychological well-being and innovative work behaviour (Mehmood et al., 2023). These findings provide a basis for investigating comparable relationships in academic workplaces rather than direct evidence about faculty. A case study of an urban elementary-school principal illustrates reflective, data-informed approaches to social-justice leadership (Ezzani, 2021). We propose that university leaders consider how comparable practices might inform their response to faculty inequity. Within this paper, manipulation is used analytically to describe interpersonal or managerial practices such as incomplete information, shifting expectations, selective consultation, or the exploitation of career dependency. Separately, research on implicit voice theories shows that employees may hold taken-for-granted beliefs that speaking up is risky or inappropriate (Detert & Edmondson, 2011). We distinguish manipulation as an interpersonal or managerial process from the broader organizational distribution of risk and opportunity examined in theories of inequality regimes and racialized organizations (Acker, 2006; Ray, 2019). This distinction is

part of the present paper's analytical framing, and the concepts are not treated as synonyms.

Academic departments distribute career-relevant resources and expectations through local structures, including workload allocation and promotion processes (O'Meara et al., 2019). Chairs and senior colleagues may influence these processes through teaching assignments, committee work, mentoring, access to resources, annual evaluation, and recommendations for reappointment or promotion (Acker, 2006; O'Meara et al., 2017). In our analytical framing, legitimate managerial authority does not itself constitute oppression; the issue is how discretion is exercised and with what consequences. Relevant concerns include unclear expectations for promotion to full professor among tenured associate professors (Kulp et al., 2019), low-recognition work is distributed inequitably (Guarino & Borden, 2017), or employees perceive speaking up as personally risky (Detert & Edmondson, 2011). Such patterns make organizational process, rather than individual intent alone, central to the analysis (Ray, 2019).

2.2. Epistemologies of Ignorance, Epistemic Exclusion, and Constrained Voice

Alcoff (2007) describes epistemologies of ignorance as processes through which not knowing is actively produced or maintained. Wilful ignorance is not simply a lack of information; it can involve refusing to engage with evidence that threatens a dominant understanding of an institution as fair (Alcoff, 2007; Pohlhaus, 2012). Pohlhaus (2012) similarly explains that dominantly situated knowers may refuse to recognize interpretive resources developed by people in marginalized positions. Applied to academic settings, these concepts suggest how faculty may know their own experiences yet encounter institutional practices that recast their accounts as subjective, unprofessional, or insufficiently objective (Alcoff, 2007; Pohlhaus, 2012).

Dotson (2011) connects ignorance with practices of silencing. Speech does not become meaningful voice merely because a person is allowed to speak; the listener must also recognize the speaker as credible and capable of contributing knowledge (Dotson, 2011). In academic settings, this problem can appear when faculty are formally invited to participate while their concerns or scholarly contributions are minimized or reinterpreted (Settles et al., 2021). Epistemic exclusion provides a higher education application of this problem (Settles et al., 2024). It identifies scholarly devaluation that can emerge when disciplinary standards and identity-based biases combine to treat particular scholars, methods, or research topics as less legitimate (Settles et al., 2021, 2024). Scholarly devaluation has been associated with lower job satisfaction, more negative perceptions of departmental climate, and stronger intentions to leave (Settles et al., 2022).

Career vulnerability can encourage silence when employees perceive that speaking up is risky or unlikely to be effective (Detert & Edmondson, 2011; Morrison & Milliken, 2000). Morrison (2014) defines employee silence as withholding concerns or work-related information that could be useful to the organization.

Employees may remain silent because of beliefs about the interpersonal or organizational risks attached to voice (Detert & Edmondson, 2011). Research identifies both uncertainty about promotion expectations among tenured associate professors (Kulp et al., 2019) and beliefs that speaking up at work is risky or inappropriate (Detert & Edmondson, 2011). We propose that uncertainty about advancement may heighten the perceived cost of challenging authority in academic settings; this relationship requires direct investigation.

2.3. Student Evaluations and Pedagogical Risk

Student evaluations of teaching are commonly used in faculty evaluation and personnel decisions, yet their interpretation as measures of teaching effectiveness remains contested (Linse, 2017; Spooren et al., 2013). Experimental and institutional studies have identified gender and cultural bias in student ratings (Fan et al., 2019; MacNell et al., 2015). Research has also identified associations between grading leniency, course difficulty, and more favourable student ratings, creating a potential incentive problem when ratings carry high stakes for faculty (Stroebe, 2020).

Responding to student feedback is not inherently problematic; student ratings can provide useful information when interpreted appropriately and in context (Linse, 2017). The organizational risk occurs when student ratings operate as a high-stakes and insufficiently contextualized measure (Spooren et al., 2013; Stroebe, 2020). Reviews of the evidence caution against treating student ratings as a direct or standalone indicator of teaching quality and emphasize careful interpretation alongside other evidence (Spooren et al., 2013). In high-stakes contexts, rigorous grading or challenging course design may therefore be perceived by faculty as creating career risk, particularly where ratings are treated as direct evidence of teaching quality (Stroebe, 2020).

Evaluation dependence can also intersect with market-oriented or student-as-consumer logics in higher education (Bunce et al., 2017; Shin & Jung, 2014). Research has documented a student-as-consumer orientation and its relationship to academic performance (Bunce et al., 2017). Cross-national research has also examined market-oriented and performance-based academic environments alongside workload and working conditions as factors associated with academic job stress (Shin & Jung, 2014). In the present survey, some respondents connected retention and tuition concerns with pressure to pass students, reduce academic challenge, or avoid conflict (see **Appendix A**). These statements are treated as reported perceptions rather than as evidence that such pressures are widespread (see **Appendix A**).

2.4. Unequal Opportunity, Service, and Workload

Academic work is commonly described as a balance among teaching, research, and service, but the allocation and recognition of these activities vary across faculty and departments (O'Meara et al., 2019). Guarino and Borden (2017) found that women faculty performed more service than men, while O'Meara et al. (2017)

demonstrated that gendered work interactions contributed to women being asked to undertake more internal service. For faculty of colour, cultural taxation can include mentoring, representation, and diversity work that benefits institutions but is not always rewarded in advancement decisions (Guillaume & Apodaca, 2022).

Unequal workload can become a career mechanism when low-recognition work reduces time for activities that carry greater weight in promotion and tenure decisions (Guarino & Borden, 2017; Guillaume & Apodaca, 2022; O'Meara et al., 2017). Faculty choices about service are also shaped by work interactions and departmental practices, including the clarity and perceived equity of workload processes (O'Meara et al., 2017, 2019). This creates a potential circular problem: faculty are expected to contribute extensively, but weakly recognized work may make it harder to satisfy the criteria used to evaluate advancement (Guarino & Borden, 2017; Guillaume & Apodaca, 2022; O'Meara et al., 2017, 2019).

The present survey included reports linking workload with institutional restructuring, vacancies, redundancies, and financial pressure (see **Appendix A**). More broadly, cross-national research identifies workload, working conditions, and performance-based management as factors associated with academic job stress (Shin & Jung, 2014). The survey accounts suggest that such pressures may be especially consequential for faculty with less authority to negotiate assignments, but this mechanism requires further investigation (see **Appendix A**).

2.5. Intersectionality and Faculty Careers

Faculty experiences of inequality are not uniform (Bonache et al., 2022; Davies & Neustifter, 2023; Johnson et al., 2025). Queer faculty have described surveillance and expectations that regulate how professionalism, gender, and sexuality may be expressed (Davies & Neustifter, 2023). Parenthood can influence academic career decisions and advancement, with women reporting greater concern about reconciling academic expectations and family responsibilities, particularly during less secure career stages (Bonache et al., 2022). Black women faculty narratives further demonstrate how invisibility, racial trauma, and the pressure of being one of few can create sustained psychological burdens (Johnson et al., 2025).

Career stage intersects with identity and appointment status (Bonache et al., 2022; Guillaume & Apodaca, 2022). Research on cultural taxation and workload equity shows that structural conditions, work expectations, gender, and race can shape faculty experiences of academic work and advancement (Guillaume & Apodaca, 2022; O'Meara et al., 2019). We therefore propose that faculty in tenure-track, adjunct, teaching-stream, and tenured appointments may experience the same department differently, and that differences in security, resources, and authority warrant direct investigation. A study of perceived oppression must distinguish recurring organizational mechanisms from the assumption that every respondent has the same level of vulnerability or experiences the same outcome (Acker, 2006; Davies & Neustifter, 2023; Johnson et al., 2025).

3. Method

3.1. Research Design

This study used an exploratory, cross-sectional online survey. Closed-response items were summarized descriptively, and four open-ended questions were examined through supplementary manifest qualitative content analysis (Elo & Kyngäs, 2008; Graneheim & Lundman, 2004). The open-text component was designed to map explicitly stated concerns and counter-experiences rather than to generate narrative accounts, establish thematic saturation, or support extensive latent interpretation (O'Cathain & Thomas, 2004). The design is therefore not described as narrative inquiry or as a fully qualitative-dominant study.

The study focused on self-reported experiences associated with teaching evaluation, treatment by colleagues, workload, mentoring, career experiences, and institutional challenges (see **Appendix B**). Because the data were collected through short survey fields rather than interviews, the analysis distinguishes the breadth of issues mentioned from the contextual depth supplied by a small number of respondents (see **Appendix B**).

3.2. Survey Instrument

The full NoviSurvey questionnaire began with an informed-consent item and included four matrix sections: six leader moral awareness items, 12 organizational ethics items, nine workplace trust items, and 11 institutional uncertainty items (see **Appendix B**). These matrix scales addressed a separate measurement model and were not analyzed in this paper (see **Appendix B**).

The present analysis used 11 demographic and professional questions and seven faculty-experience questions (see **Appendix B**). The demographic and professional items addressed age, gender, Hispanic/Latino/LatinX/Spanish origin, race, nationality, country of residence, academic role, teaching/research/service allocation, years of teaching, subject area, and institution type (see **Appendix B**). The faculty-experience section included closed questions about evaluation influence, pressure across teaching, research, and service, and differential treatment (see **Appendix B**). Four open-ended questions asked respondents to explain evaluation-related changes, describe differential treatment, characterize their academic career experience, and provide examples of institutional challenges (see **Appendix B**). **Appendix B** reproduces the analyzed questions and response formats (see **Appendix B**).

The evaluation measures differed substantively across waves and are not pooled. In the initial wave, respondents indicated whether their teaching style, grading/marking, or discipline had changed based on student evaluations (reported behavioural change). In the second wave, respondents considered whether their approach would be influenced in the hypothetical absence of student evaluations (counterfactual influence). The latter was recorded as a direct yes/no answer in the second-round workbook; the accompanying open-text explanation was used to interpret, not override, the selected response. These measures

address related concerns but cannot be treated as equivalent outcomes (see **Appendix B**).

3.3. Sampling and Data Collection

Participants were recruited through convenience sampling from February 13 to June 30, 2024. Hospitality and tourism faculty were invited through academic association listservs, including the International Council on Hotel, Restaurant, and Institutional Education, the Travel and Tourism Research Association, and the Council for Hospitality Management Education. Invitations contained a NoviSurvey link and QR code.

To broaden participation beyond hospitality and tourism, additional respondents were recruited through Amazon Mechanical Turk (MTurk). MTurk participants were required to be at least 18 years old, identify education as their employment industry, and have a Human Intelligence Task approval rate above 95%. They were compensated according to the platform protocol. The intended study population comprised faculty members with teaching and/or research appointments at postsecondary institutions, including those who also undertook academic administrative duties. Department Chairs, Program or Department Coordinators, and faculty serving in other academic leadership roles were considered eligible when they retained a faculty appointment. Accordingly, selection of "Administration" as an academic role was not, by itself, grounds for exclusion. By contrast, individuals employed exclusively in non-faculty administration or outside postsecondary education did not meet the intended faculty eligibility criterion. As the survey did not independently verify concurrent faculty appointments for respondents selecting "Administration," the role selection alone cannot establish the eligibility of all such records. Records with explicit evidence of an ineligible employment context require consistent exclusion from faculty-specific quantitative and qualitative analyses; records without sufficient information to verify status are identified as unverified rather than assumed to be ineligible. These limitations prevent retrospective confirmation of a fully verified faculty sample.

The two data files comprised 168 survey records, including 162 initial-wave and six second-wave records. The study record indicated that 182 individuals accessed the questionnaire, of whom 168 supplied survey records; because invitations were distributed by open links and the number of eligible invitees is unknown, the access-to-record proportion is not treated as a response rate. The checked screening audit identified three initial-wave records containing evidence of a non-postsecondary or otherwise incompatible employment narrative. To apply the faculty eligibility rule consistently, these three complete records were excluded from both the descriptive and open-text analyses. The resulting descriptive record set was 165 (159 first-wave, six second-wave). Retention of an administrative role alone was never an exclusion criterion; not all remaining faculty appointments were independently verified.

3.4. Data Management and Response-Level Screening

The 165 retained survey records comprise 159 first-wave and six second-wave records after respondent-level exclusions. Initial-wave data contained 67 valid Yes/No answers to the evaluation-change item and 66 valid answers each to the differential-treatment and pressure items; the six second-wave respondents answered the corresponding questions. The 92 initial-wave records without a valid evaluation answer and 93 without a valid differential-treatment or pressure answer were coded as 9/999 or otherwise unanswered. The export does not preserve a routing-path variable, so the unanswered counts cannot be reliably separated into participants not routed to the items and participants who saw but did not answer them. Accordingly, these are available-answer denominators rather than verified counts of all eligible faculty, and percentages do not describe the complete 165-record set. Missing codes 9 and 999 were excluded, rather than treated as numerical observations.

Across four open-text questions, the original files contained 243 nonmissing response units contributed by 73 distinct records: 219 answers from 67 first-wave records and 24 answers from six second-wave records. The three whole-record eligibility exclusions accounted for nine of those answers, leaving 234 candidate answers from 70 records. A second, answer-level quality screen excluded 150 further answers as nonresponsive, boilerplate or repeated, internally implausible, unsupported by a discernible faculty experience, or consisting of uninterpretable cross-references. The approved, respondent-linked screening log therefore retained 84 answers from 32 distinct contributors (62 answers from 26 first-wave records and 22 answers from six second-wave records). The log preserves anonymized record identifiers, source row, question, retention decision, and rationale. Some contributors had both retained and excluded answers; the number of qualitative response units is not treated as the number of independent respondents (see **Appendix A**).

3.5. Open-Text Depth

The revised qualitative corpus comprised 84 retained answer units from 32 distinct contributors. The median answer length was 21.5 words; 40 answers (47.6%) contained 20 words or fewer, 34 (40.5%) contained 21 - 50 words, and 10 (11.9%) exceeded 50 words. Thus, 44 answers (52.4%) exceeded 20 words, but the material remained highly uneven in contextual depth. Brief statements identified issues or clear counter-experiences without establishing organizational causes. The descriptive categories were therefore used to map manifest concerns rather than to infer prevalence or mechanisms.

A developed account was defined as a response or connected set of responses from one participant that supplied sufficient context, sequence, explanation, or reflection to support interpretation beyond the manifest issue named. Five previously identified respondents still met this threshold within the confirmed retained corpus, including two tenured respondents (see **Appendix A**). These accounts were treated as illustrative cases rather than as a representative subsample. Direct quotations were drawn from six distinct contributors, with no respondent quoted

more than once, to avoid creating an exaggerated appearance of independence.

3.6. Data Analysis

Descriptive frequencies and percentages were recalculated after removing the three records flagged by the checked respondent-level audit. The first-wave evaluation item asked about reported changes in teaching, grading/marking, or discipline associated with student evaluations (53 Yes and 14 No among 67 valid answers). The six second-wave respondents instead answered a counterfactual question about whether their approach would be influenced in the absence of evaluations (four Yes and two No). These are separate outcome measures and no combined evaluation influence percentage is calculated. The second-wave classifications were taken from the questionnaire's actual closed Yes/No responses; their open-text explanations were used for context only, not converted into or used to override binary answers. Pressure was a multiple-response item, so selected percentages do not add to 100%. No inferential tests are reported because eligibility uncertainty, convenience sampling, missingness and data-quality concerns make broader inference inappropriate.

The accompanying confirmatory factor analysis (CFA) and related statistical output were excluded from this manuscript. The output treated user-defined missing codes 9 and 999 as observed scores, producing invalid composite values, and the proposed four-factor measurement model showed unacceptable fit. Those outputs therefore cannot support the present descriptive claims. The analysis is restricted to interpretable item-level summaries and manifest open-text content.

The 84 retained answers were examined using manifest qualitative content analysis (Elo & Kyngäs, 2008; Graneheim & Lundman, 2004). The coding unit was one answer to one of the four open-ended questions; retained answers were linked by anonymized respondent ID to distinguish 84 answer units from 32 independent contributors. Descriptive coding focused on what respondents explicitly named rather than latent psychological causes. Brief affirmative and negative accounts could be retained when they conveyed interpretable responses to the relevant question; responses were not treated as evidence of structural mechanisms simply because they mentioned an issue.

Two co-authors conducted the qualitative coding and analysis. The four broad descriptive domains were informed by the survey prompts and existing literature concerning evaluation, workload, inequity, and faculty voice, while narrower indicators were organized from the explicit wording of the retained answers (see **Appendix C**). The procedure is therefore described as literature-informed manifest content analysis with response-grounded descriptive subcodes, rather than as the purely inductive development of four saturated themes. The verified answer-level screening ledger records each inclusion and exclusion decision, source question, and anonymized contributor. Related descriptive indicators were grouped into evaluation-related adaptation and constraint; workload, service, and research displacement; differential treatment, favouritism, and identity- or status-based inequity; and gatekeeping, mentoring, constrained voice, and psychological strain.

A response could contribute to more than one domain when it explicitly mentioned multiple concerns.

Although two co-authors conducted the qualitative analysis, the available study record does not establish that each answer was independently double-coded or document a formal disagreement-resolution procedure. Accordingly, no inter-coder-agreement coefficient or independently confirmed consensus procedure is claimed. The four categories are presented as auditable descriptive groupings supported by the retained-answer ledger and coding definitions, rather than as independently reliability-tested themes. The absence of a fully documented chronological codebook history and formal reflexive positionality process is acknowledged as a methodological limitation (see **Appendix C**).

4. Results

4.1. Respondent Profile and Analytic Flow

Table 1 presents the demographic and professional characteristics of the 165 records remaining after the three respondent-level exclusions. In this descriptive set, 78 respondents (47.3%) selected age 25 - 34 and 40 (24.2%) selected 35 - 44. Eighty identified as men (48.5%), 72 as women (43.6%), and two as non-binary or a third gender (1.2%); gender was missing for 11 (6.7%). In the initial wave, 43 of 159 records (27.0%) reported Hispanic/Latino/LatinX/Spanish origin. Residence strings were standardized conservatively: 130 records (78.8%) specified the United States or an identifiable US locality and 10 (6.1%) specified Canada; two entries were geographically ambiguous. These percentages describe the record set, not a representative faculty population.

The academic-role field illustrates the heterogeneity of the records. Fifty-nine respondents (35.8%) selected Administration and 17 (10.3%) selected Department Chair/Head of Subject; these roles were not excluded merely because they involved academic management responsibilities. Twenty-seven selected tenure-track positions, 11 tenured positions, nine non-tenure-track positions, 20 instructor positions and 12 adjunct/part-time positions; 10 role responses were missing (see Table 1). Teaching experience ranged from zero to more than ten years: 125 respondents reported nine years or fewer, 23 reported 10 or more years, two preferred not to answer, and 15 records were missing. The sample is not presented as an exclusively early-career cohort. Free-text subject entries frequently contained places rather than subjects and were not used to draw disciplinary comparisons.

Table 1. Demographic and professional characteristics of survey records (N = 165).

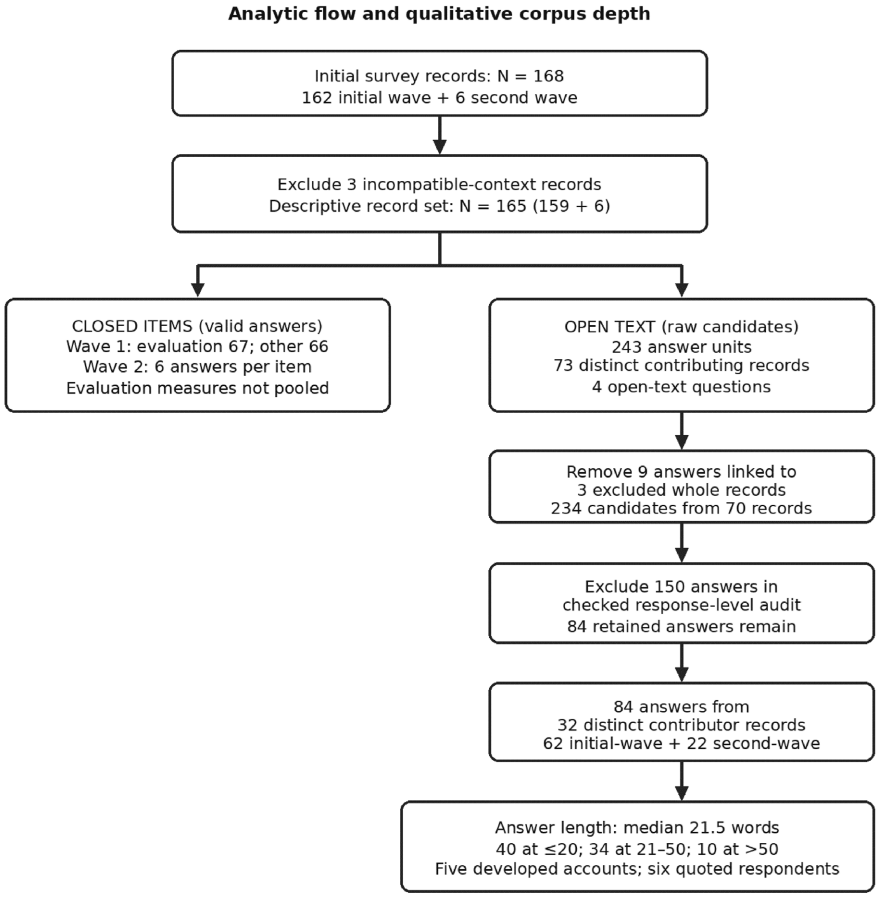
Characteristic	Category	n	%
Age	18 - 24	6	3.6
	25 - 34	78	47.3
	35 - 44	40	24.2
	45 - 54	20	12.1

Continued

	55+	9	5.5
	Prefer not to answer	1	0.6
	Missing	11	6.7
Gender	Man	80	48.5
	Woman	72	43.6
	Non-binary/third gender	2	1.2
	Missing	11	6.7
Hispanic/Latino/LatinX/Spanish origin (wave 1 only)	Yes	43	27.0
	No	105	66.0
	Missing or unrecorded	11	6.9
Residence (conservatively recoded)	United States or identifiable US locality	130	78.8
	Canada	10	6.1
	Other identifiable countries	13	7.9
	Ambiguous entry	2	1.2
	Missing	10	6.1
Current academic role	Administration	59	35.8
	Department Chair/Head of Subject	17	10.3
	Tenure-track faculty	27	16.4
	Tenured faculty	11	6.7
	Non-tenure-track faculty	9	5.5
	Instructor	20	12.1
	Adjunct/part-time instructor	12	7.3
	Missing	10	6.1
Years of teaching	0 - 2	7	4.2
	3 - 5	90	54.5
	6 - 9	28	17.0
	10 or more	23	13.9
	Prefer not to answer	2	1.2
	Missing	15	9.1
Institutional emphasis	Teaching-focused	115	69.7
	Research-focused	35	21.2
	Prefer not to answer	2	1.2
	Missing	13	7.9

Note. Three records with incompatible employment-context narratives were removed from 168 initial survey records. All percentages in this table use N = 165, except Hispanic-origin percentages (initial wave n = 159). Academic administration, including faculty chair and coordinator duties, was not an exclusion criterion. US state or locality names were conservatively standardized; two ambiguous locations were kept separate. The subject field was not reported because its free-text responses frequently named places rather than academic subjects.

Figure 1 summarizes the sequential analytic flow. The 168 original records were reduced by three record-level exclusions to 165 records. Separately, 243 candidate open-text answers were reduced by nine answers originating from the excluded records and a further 150 answer-level exclusions, leaving 84 retained answers from 32 unique contributors (see Figure 1). The closed-item analysis used question-specific valid denominators after the same record-level exclusions.



Note. Record-level eligibility screening preceded qualitative answer-level exclusions. Open-text counts are answer units, not independent people; 84 retained units came from 32 distinct records. Closed-item evaluation outcomes are separated by wave because their prompts measure different constructs.

Figure 1. Analytic flow and depth of the open-text corpus.

4.2. Closed-Item Descriptive Results

4.2.1. Influence of Student Evaluations

In the first wave, 53 of 67 valid respondents (79.1%) reported that student evaluations had changed their teaching style, grading/marking, or discipline, while 14 (20.9%) reported no such change (see Table 2). These are reports of past behaviour change among answered records.

In the second wave, four of six respondents (66.7%) answered Yes and two

(33.3%) answered No to the distinct counterfactual question asking whether removal of student evaluations would influence their approach. These are recorded binary responses, not post-hoc classifications of open-text explanations. The two waves are not pooled because their question wording and reference conditions differ (see Table 2).

4.2.2. Pressure across Teaching, Research, and Service

After respondent-level exclusions, 66 initial-wave records supplied valid answers to the multiple-response pressure item. Teaching was selected by 51 (77.3%), service by 18 (27.3%), and research by 19 (28.8%). All six second-wave respondents answered the item: two selected teaching (33.3%), five service (83.3%), and two research (33.3%). Across the comparable pressure items, teaching was selected in 53 of 72 valid records (73.6%), service in 23 (31.9%), and research in 21 (29.2%). The small second wave permits description but not meaningful between-wave comparison (see Table 2).

4.2.3. Differential Treatment

After the same record-level exclusions, 43 of 66 initial-wave respondents (65.2%) reported being treated differently from other faculty or staff, while 23 (34.8%) did not. Four of six second-wave respondents (66.7%) reported differential treatment and two (33.3%) did not. In the combined 72 valid answers to this equivalently worded question, 47 (65.3%) answered Yes and 25 (34.7%) answered No (see Table 2). The closed item does not establish whether differences were favourable, adverse, discriminatory, or systematic; interpretation depends on the details of the retained responses.

Table 2. Question-level descriptive results by data-collection wave.

Item	Response	Initial wave n/valid n (%)	Second wave n/valid n (%)	Combined (comparable items only)
Student evaluation item	Yes	53/67 (79.1)	4/6 (66.7)	Not pooled—different questions
	No	14/67 (20.9)	2/6 (33.3)	Not pooled—different questions
Pressure selected	Teaching	51/66 (77.3)	2/6 (33.3)	53/72 (73.6)
	Service	18/66 (27.3)	5/6 (83.3)	23/72 (31.9)
	Research	19/66 (28.8)	2/6 (33.3)	21/72 (29.2)
Differential treatment	Yes	43/66 (65.2)	4/6 (66.7)	47/72 (65.3)
	No	23/66 (34.8)	2/6 (33.3)	25/72 (34.7)

Note. Wave 1. evaluated reported behavioural change; wave 2 asked a counterfactual about teaching without evaluations. The second-wave Yes/No answers were recorded directly and not derived from open-text explanations, and the evaluation measures are not pooled. Item denominators exclude three incompatible-context records. Pressure was multiple response, so percentages do not total 100%.

4.3. Open-Text Corpus Characteristics

Table 3 profiles the 84 retained answer units from 32 distinct contributors. Forty responses were 20 words or fewer, 34 contained 21 - 50 words and 10 exceeded 50 words; the median was 21.5 words. These data are more selective and somewhat longer than suggested by the earlier unverified corpus description, yet sustained contextual accounts remain concentrated among five respondents. Quotations in the findings come from six distinct respondents, one per person. Accordingly, the descriptive categories indicate issues reported by selected respondents but do not represent prevalence or saturation.

Table 3. Depth profile of the open-text corpus.

Corpus characteristic	n	%	Denominator/unit
Raw nonmissing candidate answer units	243	100.0	All answers before eligibility screening
Raw contributing respondent records	73	—	Records contributing the 243 answers
Answer units removed with three excluded records	9	3.7	Of 243 initial units
Candidate answers after record exclusions	234	96.3	Of 243 initial units
Answer units excluded by answer-level screening	150	64.1	Of 234 post-eligibility candidate units
Retained qualitative answer units	84	35.9	Of 234 post-eligibility candidate units
Distinct retained contributors	32	—	Independent respondent records
Retained from initial wave	62	73.8	Of 84 retained answers
Retained from second wave	22	26.2	Of 84 retained answers
Retained answers ≤ 20 words	40	47.6	Of 84 retained answers
Retained answers 21 - 50 words	34	40.5	Of 84 retained answers
Retained answers > 50 words	10	11.9	Of 84 retained answers
Retained answers > 20 words	44	52.4	Of 84 retained answers
Median retained answer length	21.5 words	—	84 retained answers
Respondents contributing a developed account	5	—	Unique retained contributors
Tenured contributors among developed accounts	2	—	Among five developed accounts
Distinct respondents quoted in Results	6	—	One quotation per respondent

Note. One open-text answer is one response unit. The checked corpus contains 84 retained units supplied by 32 distinct respondents. Of the 84 units, 40 contained 20 words or fewer, 34 contained 21 - 50 words and 10 exceeded 50 words; the latter 10 are included in the total of 44 exceeding 20 words. Developed accounts provide contextual illustration, not representative cases.

4.4. Descriptive Categories across the Open-Text Responses

The categories in Table 4 summarize explicit concerns and counter-experiences across the four open-ended questions (see Table 4). They are descriptive categories rather than saturated interpretive themes (see Table 4). The evidence is presented at two levels: brief responses identify breadth of mention, while developed accounts provide limited illustration of possible relationships and consequences (see Table 4). Neither level is used to estimate prevalence across higher education (see Table 4).

Table 4. Descriptive categories and evidentiary boundaries.

Descriptive category	Manifest indicators in brief responses	Information-rich support	Interpretive boundary
Evaluation-related adaptation and constraint	Changes to style, examples, assignments, grading, difficulty, feedback, or misconduct decisions	Developed accounts linked evaluation concerns with tenure, retention, tuition, and academic judgment	Identifies reported influence; does not show whether changes improved learning or how often standards changed
Workload, service, and research displacement	Teaching pressure, service expectations, administration, insufficient research time, lack of resources	Developed accounts connected administrative expansion, high teaching load, restructuring, and reduced research output	Does not verify workload inequity against formal assignments or role expectations
Differential treatment, favouritism, and identity/status inequity	Sexism, favouritism, seniority, appointment status, credibility, unequal opportunities	Developed accounts illustrated gendered service, local definitions of merit, and chair discretion	Does not establish discrimination or systemic oppression in any institution
Gatekeeping, mentoring, constrained voice, and psychological strain	Lack of mentoring, changing benchmarks, bullying, stress, fear, ineffective reporting	Developed accounts described leadership transitions, unfulfilled commitments, retaliation concern, and mental-health strain	Does not establish causal processes or represent the wider corpus

Note. Categories summarize explicit content and are not mutually exclusive (see Table 4). They are not presented with frequency counts because response length and contextual detail varied substantially, and counts could imply a precision or prevalence that the corpus does not support (see Table 4). Developed accounts are illustrative rather than representative (see Table 4).

4.4.1. Evaluation-Related Adaptation and Constraint

Responses concerning evaluations ranged from constructive pedagogical adjustment to concern that ratings influenced grading, course difficulty, feedback, and academic-misconduct decisions (see Table 4). Brief answers commonly referred to changing style, adding examples, modifying assignments, or adjusting workload (see Table 4). One respondent described constructive use of feedback: "I use feedback from students to ensure that I am employing teaching activities that both engage but also surface the pragmatics of real life" (Q1).

Other responses framed evaluation influence as a constraint on professional judgment (see Table 4). A respondent explained being "conscious of how my student evaluations will be impacted ... and how important student evaluations are regarding tenure and promotion" (Q2). Developed accounts also connected evaluation pressure with student-retention and tuition concerns (see Table 4). The corpus supports the presence of these concerns, but most responses do not provide enough detail to determine how frequently evaluation influence altered academic standards or whether the changes improved or reduced learning (see Table 4).

4.4.2. Workload, Service, and Research Displacement

Workload responses explicitly mentioned teaching pressure, service expectations, administrative duties, insufficient research time, and lack of recognition (see Ta-

ble 4). Short statements established that these issues appeared across the corpus, while the developed accounts provided the clearest connections between assigned work and career consequences (see Table 4). One respondent who assumed a major administrative role early in an appointment wrote that "my administrative roles have just increased year after year, and my research output lower and lower" (Q3).

The developed accounts also described high teaching hours, repeated institutional restructuring, work outside the respondent's expertise, and service that displaced research required for advancement (see Table 4 and **Appendix A**). These accounts suggest possible relationships among workload allocation, institutional capacity, and career opportunity, but the survey format did not permit examination of workload records, formal expectations, or alternative explanations (see Table 4 and **Appendix A**).

4.4.3. Differential Treatment, Favouritism, and Identity- or Status-Based Inequity

Responses explicitly named favouritism, sexism, unequal service, differential credibility, and differences associated with appointment status or seniority (see Table 4). One respondent attributed different treatment to being "a woman" and "a newer hire" (Q4). Other brief responses referred to male faculty declining service, senior faculty having greater leeway, and access to opportunities depending on support provided to a chair (see Table 4).

The more developed accounts illustrated how local definitions of academic value could vary (see Table 4 and **Appendix A**). Research-active faculty described being criticized for insufficient attention to teaching, while other respondents described being marginalized for prioritizing undergraduate teaching and mentorship or for lacking large grants (see Table 4 and **Appendix A**). The data identify inconsistent and discretionary standards as a reported concern; they do not establish that every difference reflected discrimination or a coordinated system of oppression (see Table 4 and **Appendix A**).

4.4.4. Gatekeeping, Mentoring, Constrained Voice, and Psychological Strain

Brief responses named lack of mentoring, changing benchmarks, unclear expectations, bullying, stress, fear of asking questions, and ineffective reporting processes (see Table 4). One respondent summarized an early experience as "It was like hazing ... it was never enough" (Q5). Another stated, "I now fear to go into my office or speak out because of threats that my job is on the line" (Q6).

The developed accounts provided limited context for these statements, including unfulfilled start-up commitments, leadership transitions, concern about retaliation, and periods of mental-health treatment (see Table 4 and **Appendix A**). Positive responses also appeared: some faculty described helpful senior colleagues, enjoyment of teaching, useful feedback, and no significant challenges (see Table 4 and **Appendix A**). These counter-experiences are important because they show that mentoring, evaluation, service, and seniority were not inherently harmful; the

concern was inconsistency, opacity, or disproportionate career risk (see Table 4 and **Appendix A**).

4.5. Illustrative Information-Rich Accounts

Five respondents supplied enough connected detail to support short illustrative accounts (see Table 3 and **Appendix A**). Two of the five were tenured (see Table 3 and **Appendix A**). The accounts are summarized to show the kind of context available in the deepest material, not to imply that they represent the wider corpus (see Table 3 and **Appendix A**).

Developed Account 1 - Administrative role expansion (see **Appendix A**, Table A2). A respondent described being assigned responsibility for a graduate program in the second year of an appointment because no one else would take the role (see **Appendix A**, Table A2). The account connected increasing administration with reduced research output and learning major responsibilities without prior preparation (see **Appendix A**, Table A2).

Developed Account 2 - Isolation, bullying, and well-being (see **Appendix A**, Table A2). A respondent with an unusual appointment profile described isolation after relocating, bullying from a colleague, uncertainty about differential treatment, and a prolonged period of mental-health strain (see **Appendix A**, Table A2). The same account also described professional self-discovery and adaptation, illustrating that negative and developmental experiences could coexist (see **Appendix A**, Table A2).

Developed Account 3 - Chair pressure and retaliation concern (see **Appendix A**, Table A2). A respondent linked departmental "student as customer" language with pressure to pass students, criticism for being research active, and fear after challenging a chair (see **Appendix A**, Table A2). This account supplied context for the relationship among evaluation, local priorities, voice, and perceived career risk (see **Appendix A**, Table A2).

Developed Account 4 - Evaluation, gendered service, and advancement (see **Appendix A**, Table A2). A respondent connected concern about student evaluations with gendered service expectations, chair favouritism, insufficient research time, and the demands of reappointment and promotion (see **Appendix A**, Table A2). The account illustrated how several short-response categories could intersect within one person's experience (see **Appendix A**, Table A2).

Developed Account 5 -Institutional austerity and work intensification (see **Appendix A**, Table A2). A respondent described retention pressure, repeated redundancies, teaching outside their expertise, increased management duties, complex student needs, and neglected research (see **Appendix A**, Table A2). The account located workload concerns within institutional financial conditions rather than solely in interpersonal leadership (see **Appendix A**, Table A2).

Taken together, these five accounts deepen selected aspects of the descriptive categories, but they cannot compensate for the brevity of the remaining corpus (see Table 3 and **Appendix A**). Their analytical role is illustrative and hypothesis-

generating (see Table 3 and **Appendix A**).

5. Discussion

5.1. Interpreting Breadth without Overstating Depth

The qualitative component is an exploratory account of faculty-reported concerns, not an in-depth reconstruction of organizational processes. The checked audit retained 84 answers from 32 independent contributors, including 40 statements of 20 words or fewer. Although the median answer length was 21.5 words and 10 answers exceeded 50 words, a limited number supplied adequate context to examine sequences, competing explanations or organizational consequences. The four categories are therefore presented as descriptive groupings of what contributors explicitly mentioned; the five developed accounts offer illustrations rather than evidence of general causal mechanisms (see Tables 3-4 and **Appendix A**).

This distinction changes the status of the four categories (see Tables 3-4). They should be understood as descriptive groupings of manifest content rather than as saturated themes (see Tables 3-4). The five developed accounts demonstrate that evaluation, workload, treatment, mentoring, and voice can intersect, but they do not establish that the same relationships characterized the wider sample (see Tables 3-4). The findings are therefore best used to refine questions for interviews, institutional case studies, and analysis of workload, promotion, grievance, and evaluation records (see Tables 3-4).

5.2. Evaluation, Academic Judgment, and Institutional Pressure

The evaluation category is consistent with research showing that student ratings can reflect gender and cultural bias rather than teaching effectiveness alone (Fan et al., 2019; MacNell et al., 2015). Some respondents associated evaluations with changes to teaching-related judgment (see Table 4). These reports can be considered alongside Stroebe's (2020) analysis of incentives created by high-stakes student ratings. However, the short responses do not show whether the changes were pedagogically beneficial, strategically protective, or consequential for student outcomes (see Table 4).

Bunce et al. (2017) found an association between student consumer orientation and academic performance, while Stroebe (2020) examined incentives associated with student ratings. In the present survey, selected developed accounts connected evaluations with tenure, retention, and tuition concerns (see Table 4 and **Appendix A**). We interpret these accounts as suggesting a potential organizational mechanism for investigation, rather than as establishing that mechanism. Scholarship on student ratings emphasizes continuing questions about validity and the need for contextualized interpretation (Spooren et al., 2013). When ratings are used in high-stakes personnel decisions, they should be balanced with other evidence of teaching rather than treated as a stand-alone measure (Linse, 2017). The positive responses also show that student feedback can support constructive adaptation

when it is treated as one source of evidence rather than as a popularity measure (see Table 4).

5.3. Work Allocation, Differential Treatment, and Career Opportunity

The workload and differential-treatment categories are consistent with evidence that women may carry heavier internal service loads and that faculty from underrepresented groups may perform culturally important work that receives limited recognition (Guarino & Borden, 2017; Guillaume & Apodaca, 2022). Research also indicates that departmental conditions and practices are associated with faculty perceptions of workload equity and satisfaction (O'Meara et al., 2019). The current corpus identifies similar concerns in respondents' own words, particularly around service, research displacement, and reduced ability to refuse work (see Table 4).

The survey cannot determine whether workload differences were objectively unequal or whether they resulted from rank, role, institutional mission, individual choice, or discrimination (see Table 4 and **Appendix A**). It does, however, show that some respondents interpreted discretionary assignments and opportunity structures as favouring those with greater status, stronger relationships with leaders, or fewer constraints (see Table 4 and **Appendix A**). Future research should compare perceived workload with documented assignments, formal role expectations, and advancement outcomes (see Table 4 and **Appendix A**).

5.4. Gatekeeping, Epistemic Exclusion, and Voice

The gatekeeping and voice category aligns with research on epistemic exclusion, in which scholars or their work are devalued through formal and informal judgments (Settles et al., 2024). Respondents named changing standards, inadequate mentoring, reduced credibility, and fear of speaking (see Table 4). Separately, scholarly devaluation has been associated with lower job satisfaction, poorer perceptions of workplace climate, and stronger turnover intentions (Settles et al., 2022). We draw on Alcoff's (2007) epistemologies of ignorance and Dotson's (2011) account of silencing as theoretical resources for investigating how concerns may be discounted despite the existence of formal reporting channels.

The survey evidence is not deep enough to establish those processes in any institution (see Table 4). Nevertheless, the explicit references to fear, ineffective protection, and career consequences are consistent with research showing that employees may self-censor when they perceive speaking up as risky or ineffective (Detert & Edmondson, 2011; Morrison & Milliken, 2000). These statements support closer investigation of whether formal voice systems are experienced as credible and safe (Morrison, 2014; see Table 4).

5.5. Conditional Belonging as a Provisional Framework

Conditional belonging is proposed as a provisional mechanism linking organiza-

tional practices to the perceived terms of inclusion and career opportunity, rather than as a demonstrated theory. Inequality-regimes scholarship explains patterned distribution of resources and advantage (Acker, 2006); epistemic exclusion addresses whose scholarship and knowledge are recognized (Settles et al., 2021, 2024); and employee-silence research explains why concerns may be withheld when voice is perceived as risky (Morrison, 2014). Conditional belonging adds a specific relational proposition: where formal membership and advancement rules coexist with informal expectations of deference, availability for weakly rewarded work, and alignment with locally preferred definitions of merit, faculty may perceive that access to mentoring, recognition, discretionary opportunities, and credible voice is contingent on compliance. Resistance or nonconformity may therefore be perceived as carrying opportunity costs even when formal eligibility for those opportunities is unchanged. This proposed conformity–opportunity contingency is the framework's distinct explanatory claim, not an observed causal association in the present survey.

The developed accounts offer limited illustrations of how discretionary work allocation, evaluation pressures, perceived favouritism, and constrained voice may intersect, but they do not establish that conformity produces access or that resistance causes exclusion (see **Table 4** and **Appendix A**). The framework generates testable questions: do faculty who perceive stronger informal conformity expectations also report more conditional access to mentoring, recognition, and advancement opportunities; and is this association moderated by appointment security or departmental discretion? Interviews, longitudinal evidence, and institutional allocation records are needed to evaluate these propositions.

5.6. Methodological Implications

Open-ended survey questions can identify unanticipated issues and provide participants with space to qualify closed responses, but they often generate brief and uneven material (O'Cathain & Thomas, 2004). The present corpus demonstrates both functions (see **Table 3**). It exposed concerns not captured by the closed items, yet only a small number of respondents supplied enough detail for contextual interpretation (see **Table 3**).

Guidance on survey open questions emphasizes designing prompts to elicit the depth of data needed for the intended analysis (O'Cathain & Thomas, 2004). We recommend that future surveys request specific examples and consequences and, where feasible, offer participants an optional follow-up interview. Recruitment-source identifiers, completion time, duplicate checks, institution verification, and text-similarity screening should be built into the original protocol rather than added after data collection (Chmielewski & Kucker, 2020; Kennedy et al., 2020).

5.7. Practical Implications

The reported concerns suggest several practical priorities, although the study cannot determine their prevalence (see **Table 4**). First, institutions should use trans-

parent workload systems that record teaching, research, internal and external service, mentoring, and administrative responsibilities (O'Meara et al., 2019). Research on faculty workload also shows gender differences in service allocation, underscoring the need to examine who performs low-recognition work (Guarino & Borden, 2017). Workload records should therefore be reviewed for recurring differences by rank, employment status, gender, race, and other relevant identities, particularly given evidence of cultural taxation among faculty of colour (Guillaume & Apodaca, 2022). Service that supports inclusion and student success should also be recognized in evaluation and promotion when such work contributes materially to institutional functioning (O'Meara et al., 2017).

Second, student evaluations should not operate as the sole or dominant measure of teaching effectiveness (Spooren et al., 2013; Stroebe, 2020). Guidance on the interpretation and use of student ratings recommends contextualized use rather than treating ratings as a stand-alone measure (Linse, 2017). Institutions should combine student feedback with other evidence of teaching, such as peer review, course design, reflective teaching statements, and appropriate indicators of student learning (Spooren et al., 2013). Results should also be interpreted in light of established gender and cultural biases (Fan et al., 2019; MacNell et al., 2015).

Third, promotion, tenure, and reappointment criteria should be written clearly, communicated early, and applied consistently (Kulp et al., 2019). The present survey likewise identified changing benchmarks and unclear expectations among the concerns reported by respondents (see **Table 4** and **Appendix A**). Changes to advancement criteria and related resource commitments should therefore be documented so that expectations remain as transparent and consistent as possible (Kulp et al., 2019; see **Table 4**).

Fourth, mentoring should be structured rather than left to chance (Fountain & Newcomer, 2016). Research on faculty mentoring supports institutionally visible and sustained mentoring programs as a mechanism for teaching, research, career planning, and faculty development (Fountain & Newcomer, 2016). Multiple-mentoring approaches can broaden access to developmental relationships rather than making one mentor responsible for every form of support (de Janasz & Sullivan, 2004). Building on the multiple-mentoring approach, we recommend that faculty who are new, untenured, contingent, or preparing for advancement have access to more than one source of mentoring, including support outside the direct evaluative relationship (de Janasz & Sullivan, 2004).

Research on organizational silence explains how employees may collectively perceive speaking up as unwise (Morrison & Milliken, 2000). Research on implicit voice theories also shows that employees may self-censor when they take for granted that speaking up is risky or inappropriate (Detert & Edmondson, 2011). Drawing on this literature, we recommend credible routes for raising workload, discrimination, and retaliation concerns, with clear processes, documented follow-up, and protection from retaliation (Morrison, 2014). We also recommend

that university leadership development address recurring patterns of unequal assignments, inconsistent standards, favouritism, and retaliation. Ezzani's (2021) elementary-school case illustrates reflective and equity-focused leadership; applying comparable practices to faculty workplaces is a proposed extension that requires evaluation.

6. Limitations and Future Research

The study has important limitations. Recruitment combined academic association listservs with MTurk, and the analytic export did not identify recruitment source at respondent level. MTurk eligibility criteria captured employment in education but did not independently establish a higher education faculty appointment. Three records with incompatible employment-context narratives were excluded consistently from all analyses, but the remaining 165 records, especially administrative-role responses, cannot be treated as an externally verified roster of faculty appointments. The author-confirmed screening audit reduces obvious problems in the open-text corpus but cannot resolve potential misrepresentation elsewhere in the dataset.

The records were predominantly from the United States or identifiable US localities, with limited representation from outside North America; race and residence were self-reported and imperfectly standardized (see **Table 1**). The study cannot support global prevalence estimates or comparisons among demographic subgroups. Future studies should use purposive recruitment, faculty appointment verification, recruitment-source flags and consistent institutional and geographical coding.

Although two survey prompts referred to early-career academic experience, career stage was not defined prospectively. In the 165-record descriptive dataset, 23 respondents selected 10 or more years of teaching; within the five developed accounts, two were tenured (see **Table 1** and **Appendix A**). The results therefore cover current and retrospective accounts and are not presented as describing a single career-stage cohort.

Fourth, the evaluation question was worded differently across the two data files (see **Appendix B** and **Table 2**). The two measures are reported separately and are not treated as psychometrically or substantively equivalent (see **Appendix B** and **Table 2**). Future data collection should use identical wording and preserve recruitment-source and wave identifiers (see **Appendix B** and **Table 2**).

The principal qualitative limitation is uneven response depth and selective retention. From 243 original candidate answers, nine were removed with the three excluded records and 150 additional answers were removed through item-level screening, leaving 84 answers from 32 distinct contributors. Forty retained answers contained 20 words or fewer, while 10 exceeded 50 words; the median was 21.5 words. Five respondents contributed developed accounts, and the quotations came from six distinct people. The checked audit makes those totals reproducible but does not remove possible selection bias arising from retaining the more inter-

pretable answers. Particularly, 22 of the retained 84 answers came from the six second-wave respondents, whose open-text material was often comparatively detailed. Accordingly, neither richness of a category nor the number of answers should be interpreted as prevalence across the faculty population.

The survey did not permit follow-up questions, clarification of ambiguous statements, exploration of contradictory interpretations, or independent verification of reported events (see **Table 3**). Category descriptions should therefore not be interpreted as prevalence estimates, evidence of saturation, causal findings, or confirmation of systemic oppression (Elo & Kyngäs, 2008; Graneheim & Lundman, 2004; see **Table 3**). We recommend combining interviews or focus groups with institutional documents, workload records, promotion criteria, grievance data, and longitudinal career outcomes to investigate organizational mechanisms more directly.

The qualitative coding was conducted by two co-authors. However, the historical study record does not document their respective coding responsibilities, independent double-coding, a formal disagreement-resolution process, or a formal reflexive positionality procedure. Although the screened corpus and category boundaries are now auditable, the chronological codebook-development record is incomplete. These reporting limitations are acknowledged rather than replaced with unverified reliability or consensus claims (Berger, 2015; Elo & Kyngäs, 2008; Graneheim & Lundman, 2004).

7. Conclusion

This exploratory survey maps perceived concerns surrounding evaluation dependence, workload allocation, research displacement, differential treatment, mentoring, constrained voice and psychological strain. After record-level and response-level screening, the qualitative findings draw on 84 answers from 32 independent contributors, including five developed accounts. The evidence maps explicitly reported concerns rather than establishing their frequency among faculty, the causes of differential treatment, or recurring organizational mechanisms (see **Tables 2-4** and **Appendix A**).

The findings should therefore be understood as an exploratory mapping exercise and a foundation for deeper qualitative and institutional research (see **Tables 3-4**). Conditional belonging offers a provisional way to connect evaluation, work allocation, local definitions of merit, and constrained voice, but the framework requires testing with richer data (Acker, 2006; Morrison, 2014; Ray, 2019; Settles et al., 2021, 2024; see **Appendix A**). The practical directions identified in this paper—transparent workloads, contextualized teaching evaluation, clear advancement standards, structured mentoring, and credible routes for faculty voice—are also consistent with the broader literature on workload equity, faculty evaluation, promotion clarity, mentoring, and employee silence (Fountain & Newcomer, 2016; Kulp et al., 2019; Linse, 2017; Morrison, 2014; O'Meara et al., 2019).

Ethical Considerations

The research ethics board from Napier University, UK granted ethics approval in February 2024 (Ethics ID: 3373746).

The online cover page explained the study before respondents proceeded, and participation was voluntary. Direct identifiers were not included in the analytic files used for this revision. Quotations were selected to minimize deductive identification and are reproduced verbatim except for ellipses or bracketed text used to shorten or clarify a passage.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Acker, J. (2006). Inequality Regimes: Gender, Class, and Race in Organizations. *Gender & Society, 20,* 441-464. https://doi.org/10.1177/0891243206289499

Alcoff, L. M. (2007). Epistemologies of Ignorance: Three Types. In S. Sullivan, & N. Tuana (Eds.), *Race and Epistemologies of Ignorance* (pp. 39-57). State University of New York Press. https://doi.org/10.1515/9780791480038-003

Berger, R. (2015). Now I See It, Now I Don't: Researcher's Position and Reflexivity in Qualitative Research. *Qualitative Research, 15,* 219-234. https://doi.org/10.1177/1468794112468475

Bonache, H., Carballo, V., Chas, A., & Delgado, N. (2022). Go the Extra Mile: A Gender Approach towards Parenthood in Early-Career Academics. *Journal of Gender Studies, 31,* 125-135. https://doi.org/10.1080/09589236.2020.1836482

Bunce, L., Baird, A., & Jones, S. E. (2017). The Student-as-Consumer Approach in Higher Education and Its Effects on Academic Performance. *Studies in Higher Education, 42,* 1958-1978. https://doi.org/10.1080/03075079.2015.1127908

Chmielewski, M., & Kucker, S. C. (2020). An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results. *Social Psychological and Personality Science, 11,* 464-473. https://doi.org/10.1177/1948550619875149

Clarke, J., Hurst, C., & Tomlinson, J. (2024). Maintaining the Meritocracy Myth: A Critical Discourse Analytic Study of Leaders' Talk about Merit and Gender in Academia. *Organization Studies, 45,* 635-660. https://doi.org/10.1177/01708406241236610

Davies, A. W. J., & Neustifter, R. (2023). Heteroprofessionalism in the Academy: The Surveillance and Regulation of Queer Faculty in Higher Education. *Journal of Homosexuality, 70,* 1030-1054. https://doi.org/10.1080/00918369.2021.2013036

de Janasz, S. C., & Sullivan, S. E. (2004). Multiple Mentoring in Academe: Developing the Professorial Network. *Journal of Vocational Behavior, 64,* 263-283. https://doi.org/10.1016/j.jvb.2002.07.001

Detert, J. R., & Edmondson, A. C. (2011). Implicit Voice Theories: Taken-for-Granted Rules of Self-Censorship at Work. *Academy of Management Journal, 54,* 461-488. https://doi.org/10.5465/amj.2011.61967925

Dotson, K. (2011). Tracking Epistemic Violence, Tracking Practices of Silencing. *Hypatia, 26,* 236-257. https://doi.org/10.1111/j.1527-2001.2011.01177.x

Elo, S., & Kyngäs, H. (2008). The Qualitative Content Analysis Process. *Journal of Advanced Nursing, 62,* 107-115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Ezzani, M. (2021). A Principal's Approach to Leadership for Social Justice: Advancing Reflective and Anti-Oppressive Practices. *Journal of School Leadership, 31,* 227-247. https://doi.org/10.1177/1052684620908347

Fan, Y., Shepherd, L. J., Slavich, E., Waters, D., Stone, M., Abel, R. et al. (2019). Gender and Cultural Bias in Student Evaluations: Why Representation Matters. *PLOS ONE, 14,* e0209749. https://doi.org/10.1371/journal.pone.0209749

Fountain, J., & Newcomer, K. E. (2016). Developing and Sustaining Effective Faculty Mentoring Programs. *Journal of Public Affairs Education, 22,* 483-506. https://doi.org/10.1080/15236803.2016.12002262

Graneheim, U. H., & Lundman, B. (2004). Qualitative Content Analysis in Nursing Research: Concepts, Procedures and Measures to Achieve Trustworthiness. *Nurse Education Today, 24,* 105-112. https://doi.org/10.1016/j.nedt.2003.10.001

Guarino, C. M., & Borden, V. M. H. (2017). Faculty Service Loads and Gender: Are Women Taking Care of the Academic Family? *Research in Higher Education, 58,* 672-694. https://doi.org/10.1007/s11162-017-9454-2

Guillaume, R. O., & Apodaca, E. C. (2022). Early Career Faculty of Color and Promotion and Tenure: The Intersection of Advancement in the Academy and Cultural Taxation. *Race Ethnicity and Education, 25,* 546-563. https://doi.org/10.1080/13613324.2020.1718084

Johnson, S., Konopasky, A., & Wyatt, T. (2025). In Their Own Voices: A Critical Narrative Review of Black Women Faculty Members' First-Person Accounts of Racial Trauma across Higher Education. *Teaching and Learning in Medicine, 37,* 218-228. https://doi.org/10.1080/10401334.2024.2329680

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. G. (2020). The Shape of and Solutions to the Mturk Quality Crisis. *Political Science Research and Methods, 8,* 614-629. https://doi.org/10.1017/psrm.2020.6

Kulp, A. M., Wolf-Wendel, L. E., & Smith, D. G. (2019). The Possibility of Promotion: How Race and Gender Predict Promotion Clarity for Associate Professors. *Teachers College Record: The Voice of Scholarship in Education, 121,* 1-28. https://doi.org/10.1177/016146811912100507

Linse, A. R. (2017). Interpreting and Using Student Ratings Data: Guidance for Faculty Serving as Administrators and on Evaluation Committees. *Studies in Educational Evaluation, 54,* 94-106. https://doi.org/10.1016/j.stueduc.2016.12.004

MacNell, L., Driscoll, A., & Hunt, A. N. (2015). What's in a Name: Exposing Gender Bias in Student Ratings of Teaching. *Innovative Higher Education, 40,* 291-303. https://doi.org/10.1007/s10755-014-9313-4

Mehmood, S., Jabeen, R., Khan, M. A., Khan, M. A., Gavurova, B., & Oláh, J. (2023). Impact of Despotic Leadership and Workplace Incivility on Innovative Work Behavior of Employees: Application of Mediation-Moderation Model. *Heliyon, 9,* e19673. https://doi.org/10.1016/j.heliyon.2023.e19673

Morrison, E. W. (2014). Employee Voice and Silence. *Annual Review of Organizational Psychology and Organizational Behavior, 1,* 173-197. https://doi.org/10.1146/annurev-orgpsych-031413-091328

Morrison, E. W., & Milliken, F. J. (2000). Organizational Silence: A Barrier to Change and Development in a Pluralistic World. *The Academy of Management Review, 25,* 706-725. https://doi.org/10.2307/259200

O'Cathain, A., & Thomas, K. J. (2004). "Any Other Comments?" Open Questions on Questionnaires—A Bane or a Bonus to Research? *BMC Medical Research Methodology, 4,*

Article No. 25. https://doi.org/10.1186/1471-2288-4-25

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,* 1154-1186. https://doi.org/10.3102/0002831217716767

O'Meara, K., Lennartz, C. J., Kuvaeva, A., Jaeger, A., & Misra, J. (2019). Department Conditions and Practices Associated with Faculty Workload Satisfaction and Perceptions of Equity. *The Journal of Higher Education, 90,* 744-772. https://doi.org/10.1080/00221546.2019.1584025

Pohlhaus, G. (2012). Relational Knowing and Epistemic Injustice: Toward a Theory of *Willful Hermeneutical Ignorance*. *Hypatia, 27,* 715-735. https://doi.org/10.1111/j.1527-2001.2011.01222.x

Ray, V. (2019). A Theory of Racialized Organizations. *American Sociological Review, 84,* 26-53. https://doi.org/10.1177/0003122418822335

Settles, I. H., Jones, M. K., Buchanan, N. T., & Brassel, S. T. (2022). Epistemic Exclusion of Women Faculty and Faculty of Color: Understanding Scholar(ly) Devaluation as a Predictor of Turnover Intentions. *The Journal of Higher Education, 93,* 31-55. https://doi.org/10.1080/00221546.2021.1914494

Settles, I. H., Jones, M. K., Buchanan, N. T., & Dotson, K. (2021). Epistemic Exclusion: Scholar(ly) Devaluation That Marginalizes Faculty of Color. *Journal of Diversity in Higher Education, 14,* 493-507. https://doi.org/10.1037/dhe0000174

Settles, I. H., Jones, M. K., Buchanan, N. T., Dotson, K., Grower, P., O'Rourke, M. et al. (2024). Epistemic Exclusion: A Theory for Understanding Racism in Faculty Research Evaluations. *American Psychologist, 79,* 539-552. https://doi.org/10.1037/amp0001313

Shin, J. C., & Jung, J. (2014). Academics Job Satisfaction and Job Stress across Countries in the Changing Academic Environments. *Higher Education, 67,* 603-620. https://doi.org/10.1007/s10734-013-9668-y

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching: The State of the Art. *Review of Educational Research, 83,* 598-642. https://doi.org/10.3102/0034654313496870

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,* 276-294. https://doi.org/10.1080/01973533.2020.1756817

Appendix A

Anonymized Response Screening and Evidence Audit

The checked response-screening audit links each candidate open-text answer to an anonymous respondent ID and source question. Whole-record eligibility exclusions were applied before answer-level screening: three incompatible-context records contributed nine candidate answers, which were removed alongside those records. The remaining 234 candidate answers from 70 records were screened individually; 150 were excluded and 84 were retained from 32 independent contributors. A respondent could contribute both retained and excluded answers. The detailed ledger, including decision and reason for each answer, is retained as a separate auditable research file and not reproduced in the appendix to avoid unnecessary disclosure of response content.

Table A1. Response-level screening flow.

Stage	Decision rule	n	% of initial 243
Initial candidate corpus	Nonmissing answer in one of four prompts, before record-level eligibility screening	243	100.0
Whole-record exclusion	Nine answers belonging to three incompatible employment-context records	9	3.7
Post-eligibility candidate corpus	Initial answer units remaining after all three records removed	234	96.3
Answer-level exclusions	Nonresponse to prompt, template/boilerplate or duplicated content, implausible personal account, uninterpretable cross-reference	150	61.7
Retained qualitative corpus	Interpretable manifest answer or counter-experience, linked to anonymized individual respondent	84	34.6
Unique retained respondent records	Count each respondent ID only once across all retained answers	32	—

Note. Denominators are sequential: 243 original nonmissing answer units, minus nine from three excluded records, equals 234 eligible/uncertain-status candidate units; minus 150 answer-level exclusions equals 84 retained response units. The final corpus represents 32 distinct respondent records, not 84 independent accounts. Percentages shown in Table A1 refer to the original 243 candidate response units.

Table A2. Anonymized profiles of the five developed accounts.

ID	Primary contextual focus	Open-text areas contributing substantive detail
DA1	Administrative responsibility expanded early in appointment and was connected with declining research output	Evaluation approach; career experience; institutional challenge
DA2	Unusual appointment profile, relocation-related isolation, bullying, mental-health strain, and professional adaptation	Evaluation approach; differential treatment; career experience; institutional challenge

Continued

DA3	Student-as-customer language, criticism of research activity, conflict with chair, and fear of retaliation	Evaluation approach; differential treatment; career experience
DA4	Evaluation consequences, gendered service, chair favouritism, research displacement, and advancement pressure	Evaluation approach; differential treatment; career experience; institutional challenge
DA5	Retention crisis, redundancies, work intensification, teaching outside expertise, and neglected research	Evaluation approach; differential treatment; career experience; institutional challenge

Note. Two of the five developed-account respondents were tenured (see Table A2). Individual appointment statuses and original case identifiers are suppressed to reduce deductive identification (see Table A2). The profiles are analytical summaries. (see Table A2).

Table A3. Audit of direct quotation sources.

Quotation ID	Category illustrated	Role of quotation	Number of quotations
Q1	Evaluation-related adaptation	Constructive use of student feedback	1
Q2	Evaluation-related constraint	Perceived tenure and promotion consequences	1
Q3	Workload and research displacement	Administrative expansion and declining research	1
Q4	Differential treatment	Gender and newer-hire status	1
Q5	Gatekeeping and mentoring	Hazing and unattainable standards	1
Q6	Constrained voice and strain	Fear of speaking after perceived threats	1

Note. Each quotation ID represents a different respondent, and no respondent is quoted more than once (see Table A3). The audit prevents repeated quotation from a small number of respondents from being mistaken for broad corroboration (see Table A3).

Appendix B

Survey Items Included in the Present Analysis

The wording below is reproduced from the supplied NoviSurvey export and the second-wave data workbook (see Table B1). Original capitalization and minor typographical errors are retained where they form part of the exported instrument (see Table B1).

Table B1. Analyzed survey questions and response formats.

Item	Question wording	Response format
P1Q1	Do you agree to participate in the study?	Yes, I agree to participate; No, I do not agree to participate
P6Q1	What is your age?	18 - 24; 25 - 34; 35 - 44; 45 - 54; 55+; Prefer not to answer
P7Q1	Which gender do you most identify?	Male (including transgender man); Female (including transgender woman); Non-binary/third gender; Prefer not to answer
P8Q1	Are you of Hispanic/Latino/LatinX/Spanish origin?	Yes; No; Prefer not to answer (initial wave only)
P9Q1	How would you describe yourself?	Race categories and prefer not to answer
P10Q1	What is your nationality?	Open text
P11Q1	What country do you currently reside?	Country selection

Continued

P12Q1	What is your current role?	Administration; Department Chair/Head of Subject; Tenure Track; Tenured; Non-Tenure Track; Instructor; Adjunct/Part-time; Other
P13Q1	What percentage of your workload is teaching, research, and service?	Percentage allocation
P14Q1	How many years have you been teaching?	0 - 2 years; 3 - 5 years; 6 - 9 years; 10+ years; Prefer not to answer
P15Q1	What subject(s) do you teach?	Open text or subject selection
P16Q1	Would you identify your institution as more research or teaching focused?	Teaching; Research; Prefer not to answer
P17Q1—initial wave	Does your teaching style, grading/marking, and discipline change based on your student evaluations?	Yes; No
P17Q2—initial wave	How does your teaching style, grading/marking, and discipline change based on your student evaluations?	Open text
Second-wave evaluation Yes/No item	Would your approach to teaching style, grading, and handling academic misconduct be influenced if student evaluations were not a factor?	Yes; No (reported separately from the initial-wave item)
Second-wave evaluation explanation	How do you feel your approach to teaching style, grading, and handling academic misconduct be influenced if student evaluations were not a factor?	Open text (used for contextual interpretation, not to recode Yes/No answers)
P18Q1	Do you feel pressure to adhere more to teaching, research, or service within your institution? (Choose all that apply)	Teaching; Research; Service; multiple selections permitted
P19Q1	Do you feel you are treated differently to other faculty/staff in your department/subject group?	Yes; No
P20Q1	How are you treated differently to other faculty/staff in your department/subject group?	Open text
P22Q1	How would you describe your experience as an early career academic professional?	Open text
P23Q1	Can you provide specific examples of the challenges you have faced at your institution as an early career academic professional?	Open text

The first-wave evaluation measure concerns reported prior behavioural change; the second-wave measure concerns hypothetical influence if evaluations were absent. The measures were reported separately and no pooled evaluation percentage was calculated. The four matrix scales in the broader instrument were not analyzed in this manuscript (see **Table B1**).

Appendix C

Manifest Content Coding Framework

The coding framework is organized around four literature-informed issue areas

drawn from the survey prompts and existing scholarship. Within these domains, manifest indicators were grounded in the explicit content of retained responses. Screening was verified at the anonymized respondent-and-answer level, and a single retained response could contribute to more than one category when it explicitly raised more than one issue. The categorical descriptions are not presented as latent, independently verified themes or prevalence estimates (see Table C1).

Table C1. Coding categories, inclusion rules, and boundaries.

Category	Include when response explicitly mentions	Do not infer from	Typical manifest indicators
Evaluation-related adaptation and constraint	A change or anticipated change in teaching, grading, feedback, difficulty, assignments, or misconduct decisions connected to student evaluations	General statements about good teaching with no evaluation link; copied definitions of assessment	Style change; more examples; easier/harder grading; evaluation fear; tenure or retention concern
Workload, service, and research displacement	Teaching load, service, administration, insufficient time/resources, or displacement of research	General stress with no workload referent; ordinary job duties without perceived pressure or trade-off	Too much service; high teaching hours; administrative burden; lack of research time
Differential treatment, favouritism, and identity/status inequity	Different treatment or opportunity linked to gender, rank, appointment, seniority, research profile, or relationship with leaders	A mere yes/no answer with no description; difference that is explicitly favourable without broader inequity claim unless used as a counter-case	Sexism; favouritism; newer hire; senior leeway; opportunity tied to chair support
Gatekeeping, mentoring, constrained voice, and psychological strain	Mentoring quality, changing standards, unclear expectations, bullying, retaliation concern, inability to speak, or stated emotional/mental-health consequence	Generic career difficulty without institutional or relational context	Hazing; lack of mentor; moving benchmark; fear; bullying; stress; ineffective reporting

Note. The displayed domains were shaped by survey content and the literature; descriptive indicators derive from the retained answers. The record-level and response-level audit verifies which text units were eligible for interpretation but is not itself evidence of intercoder reliability. Brief responses can establish that a problem or counter-experience was mentioned, not its causes or severity.