

Research Article

Cite this article: Alharbi, W. (2025). Between Documentation and Pedagogy: ESL/EFL Teacher Burnout and Perceptions of AI's Potential for Workload Relief. *Educational Process: International Journal*, 17, e2025309. <https://doi.org/10.22521/edupij.2025.17.309>

Received May 5, 2025

Accepted June 20, 2025

Published Online July 8, 2025

Keywords: Teacher burnout, ESL/EFL, quality assurance, artificial intelligence, Job Demands-Resources model, Saudi higher education, feedback workload

Author for correspondence:

Wael Alharbi

✉ harbiw@rcjy.edu.sa

✉ Yanbu English Language & Preparatory Year Institute, English Department, Royal Commission for Jubail and Yanbu, Saudi Arabia

**OPEN ACCESS**

© The Author(s), 2025. This is an Open Access article, distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction, provided the original article is properly cited.

Between Documentation and Pedagogy: ESL/EFL Teacher Burnout and Perceptions of AI's Potential for Workload Relief

Wael Alharbi 

Abstract

Background/purpose. ESL/EFL teachers face unique burnout challenges due to quality assurance (QA) requirements, increasing class sizes, and intensive feedback demands. This research investigates how these factors contribute to teacher burnout in Saudi higher education and explores the potential of AI as a workload solution. Grounded in the Job Demands-Resources model and Technology Acceptance Model, the study examines burnout drivers, feedback practices, class size effects, AI perceptions, and demographic variation.

Materials/methods. This mixed-methods study collected quantitative data from 258 ESL/EFL teachers and qualitative insights from 15 interviews and 5 focus groups. Analysis employed structural equation modeling, LASSO regression, and mediation analysis to examine burnout mechanisms and potential technological interventions.

Results. Findings reveal that QA standards significantly increase workload while reducing teachers' ability to provide individualized feedback, with workload acting as a powerful mediator of teacher wellbeing. While educators recognize AI's potential to support administrative tasks, significant adoption gaps persist—particularly among highly qualified and experienced staff who express concerns about professional identity and pedagogical displacement.

Conclusion. This study uniquely frames AI as a classroom assistant rather than a replacement for pedagogical judgment, offering empirical evidence that teacher-centered AI integration could alleviate workload stress while preserving professional autonomy. Recommendations include recalibrating QA implementation and developing AI systems that complement rather than replace teacher expertise in ESL/EFL contexts.

1. Introduction

English as a Second/Foreign Language (ESL/EFL) education faces unique challenges that distinguish it from content-focused disciplines. Unlike subjects centered primarily on knowledge transmission, language teaching remains fundamentally skills-oriented, requiring continuous cycles of practice, feedback, and personalized guidance. This pedagogical imperative creates distinctive workload patterns for ESL/EFL educators, who must provide detailed feedback across multiple linguistic dimensions while managing increasingly larger classes. Simultaneously, the expansion of quality assurance (QA) frameworks has introduced substantial administrative requirements that compete for teachers' limited time and cognitive resources. This confluence of pressures—delivering personalized instruction to growing student numbers while completing extensive documentation—has created fertile conditions for teacher burnout throughout ESL/EFL contexts globally.

These pressures create a vicious cycle where insufficient feedback time leads to deteriorating educational outcomes while institutional demands intensify (Chen et al., 2024; Ding, 2024). The emotional toll on teachers—caught between professional commitment and overwhelming workload realities—compounds this situation, as educators experience profound guilt regarding their inability to provide adequate feedback within existing institutional constraints (Ma & Liu, 2024; Sadoughi et al., 2024).

Despite extensive research on teacher burnout and emerging interest in educational AI applications, three critical gaps remain: (1) limited understanding of how language teaching's skill-based nature creates unique vulnerability to workload intensification; (2) insufficient examination of how quality assurance frameworks specifically impact ESL/EFL teachers' wellbeing; and (3) lack of empirical investigation into how AI might address these challenges while preserving pedagogical quality and professional identity. While previous studies have documented general burnout factors (Maslach & Jackson, 1981; Siyum, 2023), they have not adequately addressed the distinctive intersection of feedback demands and administrative requirements that characterizes ESL/EFL contexts. This study addresses the critical gap in understanding how AI might function as a classroom assistant to alleviate systemic burnout drivers in ESL/EFL contexts while preserving professional autonomy and pedagogical quality.

Amid these converging challenges and accumulating pressures, the advent of revolutionary artificial intelligence technology may offer help by alleviating certain aspects of teacher workload. Beyond simple automated grammar checks and surface-level flagging of writing errors, emerging AI tools have proven extremely powerful in analyzing language learner texts and suggesting tailor-made improvements. This may potentially reduce time spent on highlighting mechanical feedback, allowing teachers to concentrate on spending more time with their students. Nevertheless, several questions remain unanswered regarding language teachers' attitudes toward this technology, their knowledge of its potential, implementation roadblocks, and its actual impact on their workload and wellbeing in real-life educational settings. Striking the right balance between technological assistance and essential human judgment remains an important consideration when exploring sustainable solutions to teacher burnout.

This study explores the complex interrelationships between academic quality assurance requirements, classroom size, assessment practices, and teacher wellbeing outcomes in ESL/EFL teaching environments. Uniquely, it examines the potential role of AI when positioned not as a pedagogical disruptor, but as a teacher assistant that alleviates administrative strain and enhances feedback workflows. By investigating these relationships across different teacher demographics and institutional contexts, we aim to identify targeted intervention strategies that could support teacher sustainability while maintaining teaching quality.

Our research is theoretically grounded in the Job Demands-Resources model (Bakker & Demerouti, 2017), which conceptualizes burnout as resulting from imbalances between job demands and available resources. We extend this framework by examining how quality assurance frameworks simultaneously function as both a demand (increasing documentation requirements) and a theoretical resource (supporting instructional quality). This paradoxical relationship, combined with insights from Technology Acceptance Model (Davis, 1989) and Self-Determination Theory (Ryan & Deci, 2023), provides a robust theoretical foundation for understanding both burnout mechanisms and potential technological interventions.

Through a mixed-methods approach combining statistical modeling and qualitative insights, this study provides a nuanced perspective on how burnout emerges in ESL/EFL contexts and how institutions might respond. By gathering data on teacher experiences, perceptions of AI tools, and institutional factors, we develop an integrated understanding of both the challenges facing language educators and the potential pathways toward more sustainable teaching practices that preserve pedagogical quality and professional identity.

2. Literature Review

This literature review synthesizes four interconnected research domains: (1) the evolving landscape of teacher burnout with particular attention to ESL/EFL contexts; (2) workload intensification and accountability frameworks; (3) feedback practices and challenges in language instruction; and (4) emerging AI applications and their potential for alleviating teacher burnout. Throughout, we examine how theoretical frameworks including the Job Demands-Resources model, Conservation of Resources theory, Technology Acceptance Model, and Self-Determination Theory illuminate these phenomena and provide a foundation for our study.

Teacher burnout has been growing as a serious global issue (Agyapong et al., 2023) affecting teaching quality, student performance, and workforce stability (An & Tao, 2024). An and Tao's (2024) study of 245 EFL teachers identified significant negative correlations between burnout and both instructional effectiveness and student achievement. Burnout manifests as emotional exhaustion, depersonalization, and reduced personal accomplishment (Maslach & Jackson, 1981; Derakhshan et al., 2022). Pressley's research during COVID-19 revealed significant teacher burnout across multiple studies. His survey of 388 K-12 teachers found that various factors contributed to increased burnout during the pandemic (Pressley, 2021a), while a separate study of 329 elementary teachers documented elevated stress and anxiety levels (Pressley, 2021b). These findings indicated that burnout symptoms strongly predicted intention to leave the profession and instructional disengagement, with approximately 40-60% of higher education teachers experiencing burnout (Manuel et al., 2018). The COVID-19 pandemic exacerbated this situation, with emotional exhaustion among educators rising by 17% from 2019 to 2022 (Pressley & Ha, 2022).

Foreign language educators face distinctive burnout risks beyond those of other disciplines (Talbot & Mercer, 2018). Their cross-cultural study documented significantly higher emotional exhaustion among language teachers compared to other subject instructors, identifying "linguistic performance pressure" as a unique stressor. ESL/EFL instruction demands continuous cycles of practice, feedback, and personalized attention (Hyland & Hyland, 2019; Yu et al., 2021). Non-native English-speaking teachers face the additional burden of maintaining their own language proficiency, termed "linguistic vigilance overload" (Seo, 2023), while native-speaking instructors experience "identity strain" when navigating intercultural dynamics (Pennington & Richards, 2016; Zarrinabadi et al., 2023). Zarrinabadi et al.'s (2023) study with 187 EFL teachers showed how linguistic authority issues created significant emotional labor predicting burnout, particularly in diverse contexts. These challenges are compounded by the rise of "audit culture," transforming educational governance into a performance-driven system (Sun & Zhang, 2023).

Workload intensification is a primary driver of teacher burnout across diverse educational systems (Wang, 2022; Chang, 2022; An & Tao, 2024; Sadoughi et al., 2024; Nayernia & Babayan, 2019; Droogenbroeck et al., 2014; Chen et al., 2024; Ma & Liu, 2024; Ding, 2024). Chang's (2022) study of 319 EFL teachers identified workload intensity as the strongest predictor of emotional exhaustion, accounting for 32% of burnout variance. Educators work 54-59 hours weekly during term time, with up to 40% on non-instructional tasks (Walker et al., 2022). In higher education, quality assurance (QA) frameworks significantly contribute to this intensification, with institutions managing 14-22 distinct reporting systems (Ma & Liu, 2024). Their institutional ethnography documented how QA requirements created "bureaucratic displacement"—diverting teacher attention from pedagogical activities to administrative compliance, creating what Schaufeli et al. (2023) term the "quality paradox," where mechanisms designed to improve standards inadvertently undermine effectiveness.

In ESL/EFL contexts, teachers spend 8.7-14.2 hours weekly on QA documentation (Almossa & Alzahrani, 2022), consuming approximately 22% of total working hours, significantly higher than the 17% in other disciplines. This "documentation density" reduces time for lesson planning and correlates strongly with emotional exhaustion (Kovalkova & Malkova, 2021). The conflict between administrative demands and teaching priorities creates what Kenny and Fluck (2022) describe as the "quality-autonomy dilemma," where accountability measures erode teachers' autonomy and job satisfaction through "procedural confinement."

Effective language acquisition relies heavily on detailed written corrective feedback (WCF) as a cornerstone of ESL/EFL instruction (Ferris, 2022). Meta-analysis shows WCF significantly enhances learning outcomes across 42 studies (Li & Vuono, 2019). However, providing quality feedback requires attention to multiple linguistic dimensions (Zhang & Hyland, 2018) through four cognitive phases: error identification, linguistic analysis, feedback formulation, and learning pathway consideration. Teachers spend 20-40 minutes per paper, translating to 8-12 hours per assignment cycle for typical classes (Rajab et al., 2016).

Increasing class sizes exacerbate these challenges. In Saudi universities, the average EFL class size has risen from 22 students in 2016 to 38 in 2023 (Alutaybi & Alfares, 2024), with the most dramatic increases (73%) in foundation programs. Research establishes 25 students as the threshold beyond which individualized feedback becomes compromised (Zhang & Hyland, 2022; Barrot, 2023; Shi & Aryadoust, 2024; Ko, 2022). Shadiev and Feng's (2023) analysis found that classes exceeding 25 students showed a 47% reduction in feedback specificity and a 62% decrease in metalinguistic explanation. This "time-consuming feedback" (Alshahrani & Storch, 2025) forces teachers to choose between providing superficial feedback to all students or detailed guidance to a few.

The emotional labor of providing feedback significantly contributes to burnout (Bodenheimer & Shuster, 2019; Yin et al., 2019; Lee, 2017; Pike et al., 2019). Bodenheimer and Shuster's (2019) study revealed that balancing critical feedback with student motivation represents "assessment emotional labor," requiring teachers to suppress their reactions while crafting balanced responses, significantly predicting emotional exhaustion. Administrative demands restrict time for meaningful student interactions, with teachers reporting an average reduction of 6.4 hours weekly in student interaction following new documentation mandates (Ma & Liu, 2024).

These converging pressures create a "pedagogical sustainability crisis" manifesting as "motivation-practice incongruence" (Schaufeli et al., 2023), where teachers' professional commitment conflicts with practical limitations. This incongruence is a potent predictor of burnout compared to workload alone, creating cognitive dissonance about professional purpose and shifting burnout framing from individual deficiency to a structural problem requiring systemic solutions.

Recent AI developments offer possibilities for reducing teacher workload (Zhou et al., 2023; Jonäll, 2024). Alsalem (2024) found basic AI tools reduced grading time by approximately 40%, while Steiss et al. (2024) demonstrated that AI-generated feedback can match and, in some cases, exceed human feedback quality, with students receiving AI-supported feedback showing greater improvement in writing compared to those receiving traditional feedback. Alsalem's quasi-experimental study found teachers reallocated 4.7 hours weekly to other instructional activities after AI adoption, with students reporting higher satisfaction with feedback timeliness and helpfulness.

More sophisticated systems, such as ChatGPT, are capable of performing complex feedback tasks. In a comparative study, Steiss et al. (2024) demonstrated that AI-generated feedback can match and, in some cases, exceed human feedback in key areas. Their findings identified five core functions of AI-augmented feedback: error identification, linguistic explanation, revision suggestion, resource recommendation, and progress tracking. Notably, students who received AI-supported feedback showed greater improvement in their writing, especially among intermediate-level learners, compared to those who received traditional feedback.

However, AI adoption faces challenges. Guo and Wang (2023) identified the "AI Trust Paradox" where teachers recognize time-saving potential while fearing professional devaluation. Their survey of 412 language teachers revealed 78% acknowledged potential benefits while 63% expressed professional role concerns. Adoption intentions correlated with perceived usefulness and ease of use, while negatively associated with identity threat concerns.

Ethical considerations further complicate implementation. Dakakni and Safa's (2023) analysis uncovered algorithmic bias, with feedback to non-native writers containing more prescriptive language and fewer acknowledgments of stylistic variation, demonstrating "algorithmic linguistic imperialism." Alghamdy's (2023) legal analysis identified significant regulatory gaps regarding algorithmic transparency and bias mitigation.

Research suggests the "teacher-in-the-loop" approach (Colonna, 2023), where AI functions as an assistant rather than a replacement, is optimal (Chen et al., 2023). Colonna's (2023) action research showed this approach yielded similar workload benefits to fully automated systems (35% vs. 39% time reduction) but generated significantly higher teacher satisfaction (mean difference = 18.7 points) and enhanced professional agency. Comprehensive training covering practical applications and ethical dimensions is essential for responsible technology use (Malik et al., 2024).

The Job Demands-Resources model (Bakker & Demerouti, 2017) conceptualizes burnout as resulting from imbalances between job demands and resources, with particularly strong effects in education. Conservation of Resources theory (Hobfoll et al., 2018) explains how resource depletion triggers downward spirals. The Technology Acceptance Model (Davis, 1989; Persico et al., 2014; Zhou et al., 2022; Naidoo, 2022) helps understand factors affecting integration. Self-Determination Theory (Ryan & Deci, 2023) emphasizes autonomy, competence, and relatedness for motivation. Administrative intensification undermines autonomy satisfaction, the strongest predictor of teaching motivation. Technologies enhancing professional autonomy correlate negatively with burnout symptoms (Sato et al., 2022), suggesting that the implementation approach significantly affects well-being (Pishghadam et al., 2023).

While literature explores ESL/EFL burnout drivers, it remains underdeveloped in positioning AI as a practical assistant. Emerging research supports AI's potential (Alsalem, 2024; Praphan & Praphan, 2023; Zhou & Hou, 2025), yet concerns remain (Guo & Wang, 2023; Dakakni & Safa, 2023). However, a persistent gap exists between demonstrated AI effectiveness and sustainable classroom integration. This study builds on the "teacher-in-the-loop" framework (Colonna, 2023), addressing professional identity concerns identified as primary adoption barriers.

This literature review has identified a critical gap in understanding how AI might address systemic drivers of ESL/EFL teacher burnout. To address this gap, we propose a conceptual model (Figure 1) integrating research on burnout, workload intensification, and quality assurance pressures, positioning AI as a potential classroom assistant rather than a replacement for teacher expertise. This model synthesizes Bakker and Demerouti's (2017) Job Demands-Resources theory with Colonna's (2023) teacher-in-the-loop approach.

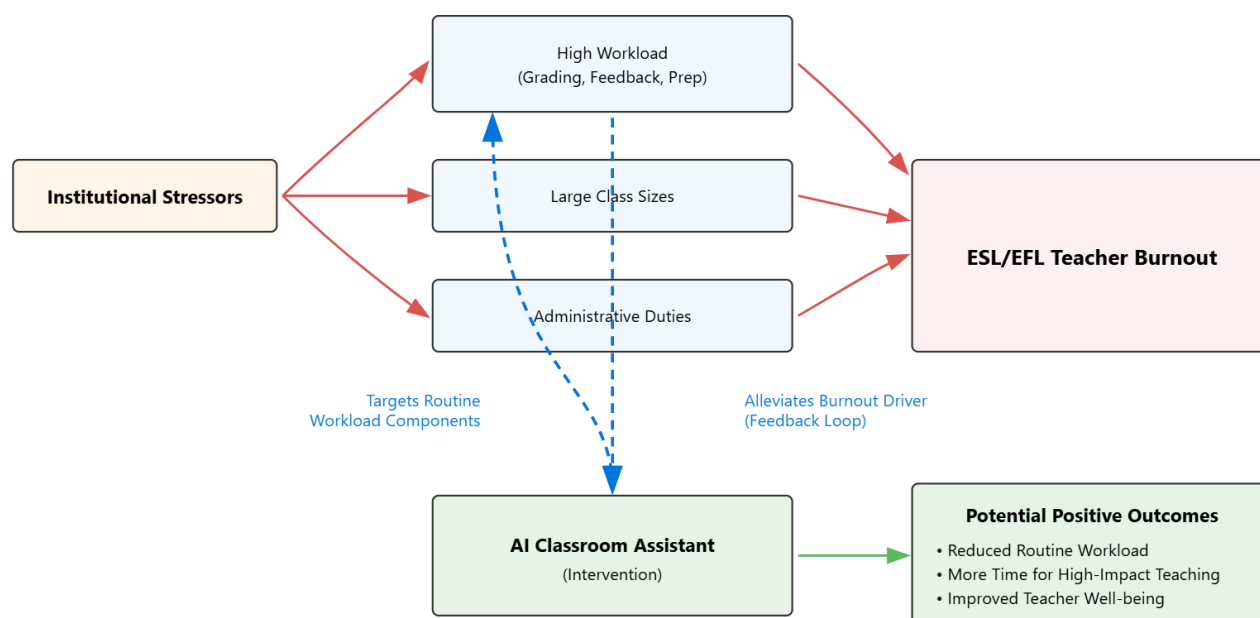


Figure 1. AI as a Classroom Assistant in EFL Contexts: A Conceptual Model

**Figure 1 shows the conceptual model illustrating systemic burnout drivers in ESL/EFL higher education and the proposed role of AI as a classroom assistant. Red arrows indicate institutional stressors contributing to teacher burnout. Green arrows represent potential positive outcomes of AI-supported feedback. The dashed blue arrow indicates a conceptual feedback loop where AI intervention could help alleviate burnout by targeting routine workload components.*

Our model identifies three primary sources of job demands creating unsustainable conditions for ESL/EFL teachers: (1) quality assurance requirements, which increase administrative workload; (2) growing class sizes, which multiply feedback demands; and (3) the inherent complexity of language feedback, which requires substantial time and emotional labor. These demands interact to create what we term 'pedagogical overload' a state where teaching responsibilities systematically exceed available time and resources. The distinctive contribution of our framework is its positioning of AI as a classroom assistant that could help rebalance the demands-resources equation (Colonna, 2023; Alsalem, 2024; Zhou et al., 2023).

Guided by this conceptual framework and the identified gaps in understanding, our study addresses the following research questions (RQs):

RQ1: What are the primary factors contributing to burnout among ESL/EFL teachers?

RQ2: How do quality assurance standards affect teacher workload and well-being?

RQ3: What challenges do teachers face in providing feedback, and how do these contribute to burnout?

RQ4: How do teachers perceive AI tools, and what potential do these tools have for reducing burnout?

RQ5: What demographic factors influence teacher burnout and openness to AI solutions?

3. Methodology

This study employed a mixed-methods design investigating relationships between teacher burnout, workload, quality assurance standards, feedback practices, and AI tools' potential in ESL/EFL contexts.

3.1. Participants

The study surveyed 258 ESL/EFL higher education teachers (Table 1) with 58% male, 42% female; 72% under 40 years; 50% doctoral degrees, 15% master's degrees, 36% bachelor's degrees; experience ranging from <5 to >21 years; 49% teaching classes of >26 students.

Table 1. Participant Demographics

Characteristic	Distribution
Gender	58% male, 42% female
Age	72% under 40 years
Qualifications	50% doctoral, 15% master's, 36% bachelor's degrees
Teaching experience	Range from <5 to >21 years
Class size	49% teaching classes of >26 students

3.1.1. Study Population and Sampling Strategy

This study employed a purposive geographic sampling approach, selecting one higher education institution from each major geographic region of Saudi Arabia (Northern, Eastern, Western, Central, and Southern regions). Rather than targeting the entire population of ESL/EFL teachers nationwide, we focused on representative institutions to capture regional diversity while maintaining methodological feasibility.

Sampling frame: The target population comprised ESL/EFL teachers employed at the selected higher education institutions across the five geographic regions (N ≈ 457 based on institutional records from participating universities).

Selection procedure: At each participating institution, we utilized institutional gatekeepers - personal contacts within the ESL/EFL departments who served as intermediaries to distribute the study materials. These contacts were asked to pass the participation message, electronic consent form, and questionnaire link to their ESL/EFL colleagues within their respective institutions. This gatekeeper-mediated convenience sampling approach was necessary due to institutional access constraints and privacy considerations.

Inclusion criteria required: (1) current employment as ESL/EFL instructor at one of the participating institutions, (2) minimum one year teaching experience, and (3) willingness to participate in potential follow-up interviews.

Sample characteristics: The final sample of 258 teachers was recruited through these institutional contacts, achieving a response rate of 64.5% from approximately 400 teachers who received the invitation through their colleagues across the five participating institutions.

Representativeness considerations: While this convenience sampling approach through gatekeepers may introduce selection bias, geographic diversity was ensured through: (1) inclusion of institutions from all five major regions, (2) representation of different institutional contexts within each region, and (3) diverse participant demographics across career stages and qualifications. The sample comprised 58% male, 42% female teachers; 72% under 40 years; 50% doctoral degrees, 15%

master's degrees, 36% bachelor's degrees; experience ranging from <5 to >21 years; 49% teaching classes of >26 students.

Following the survey, 15 individual interviews and five focus groups (4-6 participants each) were conducted with purposively selected participants representing different burnout profiles: high burnout/high workload (n=5), high burnout/moderate workload (n=3), moderate burnout/high workload (n=4), and low burnout (n=3). Focus groups were organized thematically (early career teachers, experienced teachers, large classes, AI perspectives, and quality assurance), with participants selected to represent diverse attitudes toward AI and quality assurance.

3.2. Instruments and Procedures

The survey comprised six constructs: (1) Background Information; (2) Work-Related Activities; (3) Teacher Workload and Stress; (4) Quality Assurance Standards; (5) Written Corrective Feedback; and (6) AI-Assisted Tools. All instruments underwent piloting with 28 ESL/EFL teachers through cognitive interviews (Willis, 2005), resulting in modifications to improve clarity and reliability. Low-performing items ($\alpha < 0.70$) were revised or eliminated, reducing completion time from 35 to 25 minutes. Interview and focus group protocols were tested with six participants, leading to refinements in question sequencing and additional prompts (Appendix III).

Surveys were administered online over six weeks. Interviews (45-60 minutes) and focus groups (90 minutes) were conducted via video conferencing, recorded, and transcribed. All qualitative data collection occurred within two months of survey completion.

3.3. Data Analysis

Quantitative data were analyzed using R (version 4.3.0) through multiple complementary techniques, each selected for specific theoretical and methodological reasons (Field, 2018; Tabachnick & Fidell, 2019):

Non-parametric testing justification: Despite using regression analysis, non-parametric tests (Kruskal-Wallis and Wilcoxon Signed Rank tests) were conducted because: (1) Shapiro-Wilk tests revealed significant deviations from normality in key variables ($W < 0.95$, $p < 0.001$), (2) Likert scale data may not meet interval-level assumptions required for parametric tests (Jamieson, 2004), and (3) non-parametric tests provide robust alternatives that make fewer distributional assumptions (Hollander et al., 2014).

LASSO regression justification: LASSO (Least Absolute Shrinkage and Selection Operator) was selected over conventional stepwise regression for several methodological advantages (Tibshirani, 1996; Hastie et al., 2009): (1) LASSO handles multicollinearity more effectively by shrinking coefficients toward zero rather than making binary inclusion/exclusion decisions, (2) it provides better prediction accuracy when many predictors are correlated, as confirmed by our correlation matrix showing extensive collinearity, (3) LASSO reduces overfitting through regularization, particularly important with our sample size relative to predictor variables, and (4) it performs automatic variable selection while maintaining model parsimony (James et al., 2013). Stepwise regression, while interpretable, suffers from instability in variable selection and can produce models that fail to replicate (Flom & Cassell, 2007).

The analytical sequence included: descriptive statistics, reliability analysis (Cronbach's alpha), correlation analysis, principal component analysis for dimensionality reduction, LASSO regression for predictor identification, structural equation modeling for causal pathway testing, and mediation analysis for examining indirect effects.

LASSO regression addressed multicollinearity while identifying influential variables; structural equation modeling tested causal pathways; and mediation analysis examined how AI perceptions potentially mediate relationships between workload and burnout.

Qualitative data underwent thematic analysis using NVivo (version 12) following Braun and Clarke's (2006) framework. Two researchers independently conducted initial coding (84% inter-rater agreement), followed by focused coding and theme development through iterative categorization.

3.4. Integration and Ethical Considerations

This study employed an explanatory sequential mixed methods design (Creswell & Clark, 2017), specifically following the QUAN → qual model where quantitative data collection and analysis occur first, followed by qualitative data collection designed to explain and elaborate on quantitative findings.

Design rationale: The explanatory sequential approach was selected because: (1) our primary research questions required quantitative analysis to identify relationships between variables, (2) qualitative data was needed to explain mechanisms underlying statistical relationships, and (3) this design allows for purposive sampling of qualitative participants based on quantitative results (Tashakkori & Teddlie, 2010).

Implementation sequence:

Phase 1 (QUAN): Survey administration (n=258) → statistical analysis → identification of key patterns and relationships

Interface: Selection of interview/focus group participants based on survey responses and burnout profiles

Phase 2 (qual): Individual interviews (n=15) and focus groups (n=5) → thematic analysis

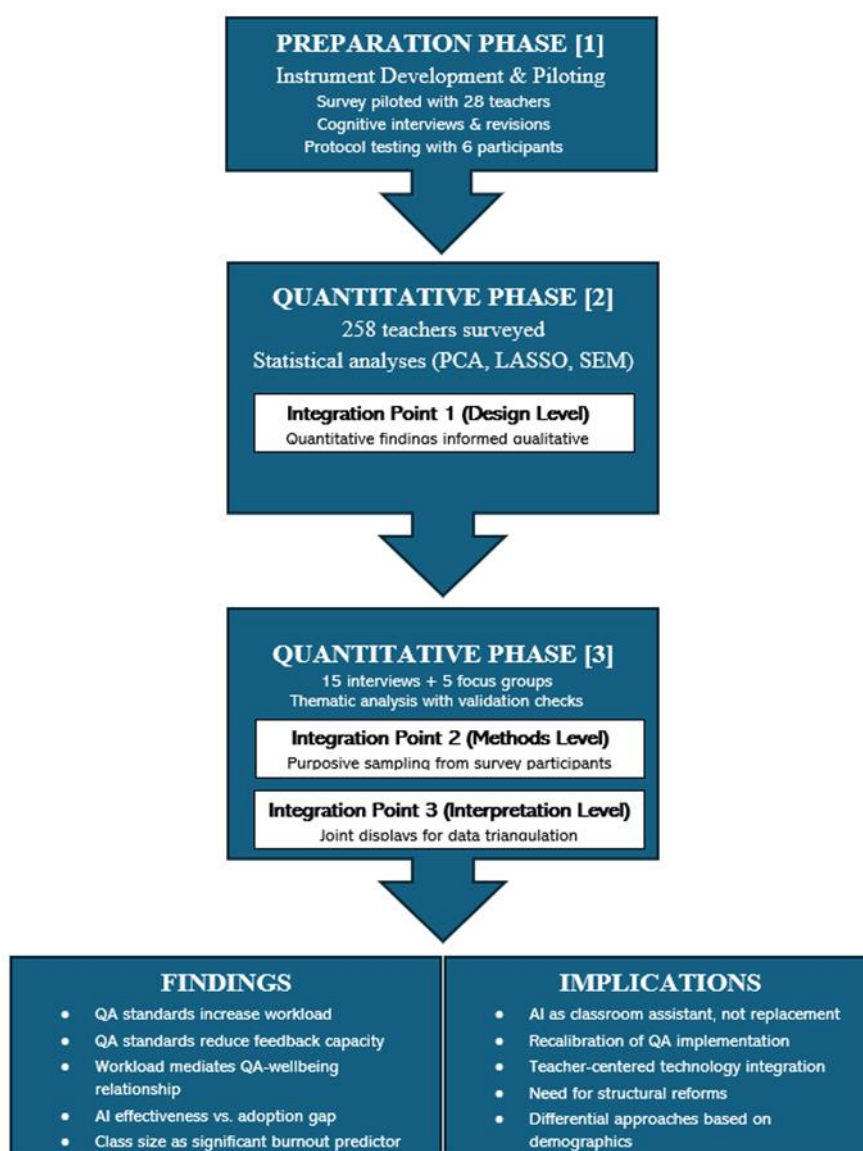
Integration: Joint interpretation of quantitative patterns and qualitative explanations

Priority weighting: QUAN emphasis (70%) with qual elaboration (30%), reflecting our focus on measuring relationships while using qualitative insights to explain quantitative findings (Creswell & Clark, 2017). Validity was strengthened through member checking, peer debriefing, and an audit trail (Fetters et al., 2013).

The integration process connected statistical relationships with qualitative themes. For example, when SEM revealed a significant relationship between quality assurance standards and reduced feedback capacity ($\beta = -0.224$, $p < 0.001$), qualitative data identified the 'documentation density' theme explaining how administrative requirements displaced feedback time. Similarly, correlations between class size and emotional exhaustion were contextualized through focus group narratives, revealing specific threshold points for feedback quality deterioration.

The study adhered to ethical principles, including informed consent, confidentiality, and voluntary participation. Limitations include the cross-sectional design limiting causal inferences, reliance on self-reported measures, generalizability concerns, and potential impacts of virtual data collection.

Figure 2. Mixed-Methods Research Design showing instrument validation, data collection phases, and integration points in the explanatory sequential design



4. Results

We analyzed quantitative data using descriptive statistics, correlation analyses, factor analysis suitability (Kaiser-Meyer-Olkin), Cronbach's alpha for internal consistency, Kruskal-Wallis tests for demographic differences, variable inflation factors for collinearity assessment, principal component analysis (PCA) for dimensionality reduction, LASSO regression for predictor identification, structural equation modeling for relationship testing, and Wilcoxon Signed Rank Tests for comparing perceptions. All analyses were conducted using R version 4.3.0 (R Core Team, 2023).

4.1. Demographic Characteristics and Descriptive Statistics

Survey respondents (n=258) were predominantly male (58% vs. 42% female) and younger (72% under age 40), as shown in Table 2. Half held doctoral degrees, with only 15% qualified at the master's level. Class sizes showed polarization, with 49% teaching large classes (>30 students) and 36% teaching smaller classes (10-15 students).

Table 1. Breakdown of demographic variables (n=258)

Variable	Factor level	Count	Percentage*
Gender	Male	149	58%
	Female	109	42%
Age Group	22 - 29	92	36%
	30 - 39	93	36%
	40 - 49	37	14%
	50 - 59	18	7%
	> 60	18	7%
	BA degree	92	36%
	MA degree	38	15%
Highest Qualification	Doctorate	128	50%
	Other	0	0%
Teaching Years	0 - 5	92	36%
	6 - 10	93	36%
	11 - 15	37	14%
	16 - 20	18	7%
	> 21	18	7%
	10 - 15	92	36%
	16 - 20	19	7%
Class Size	21 - 25	20	8%
	> 26	127	49%

* Percentages are rounded to the nearest whole percent and therefore may not total 100.

Descriptive statistics revealed a mean response of 2.4 across all questions. Response distributions (Figure 3) confirmed that most questions elicited polarized responses with few neutral selections, suggesting teachers hold strong opinions about workload issues rather than ambivalence—underscoring the emotionally charged nature of these professional challenges.

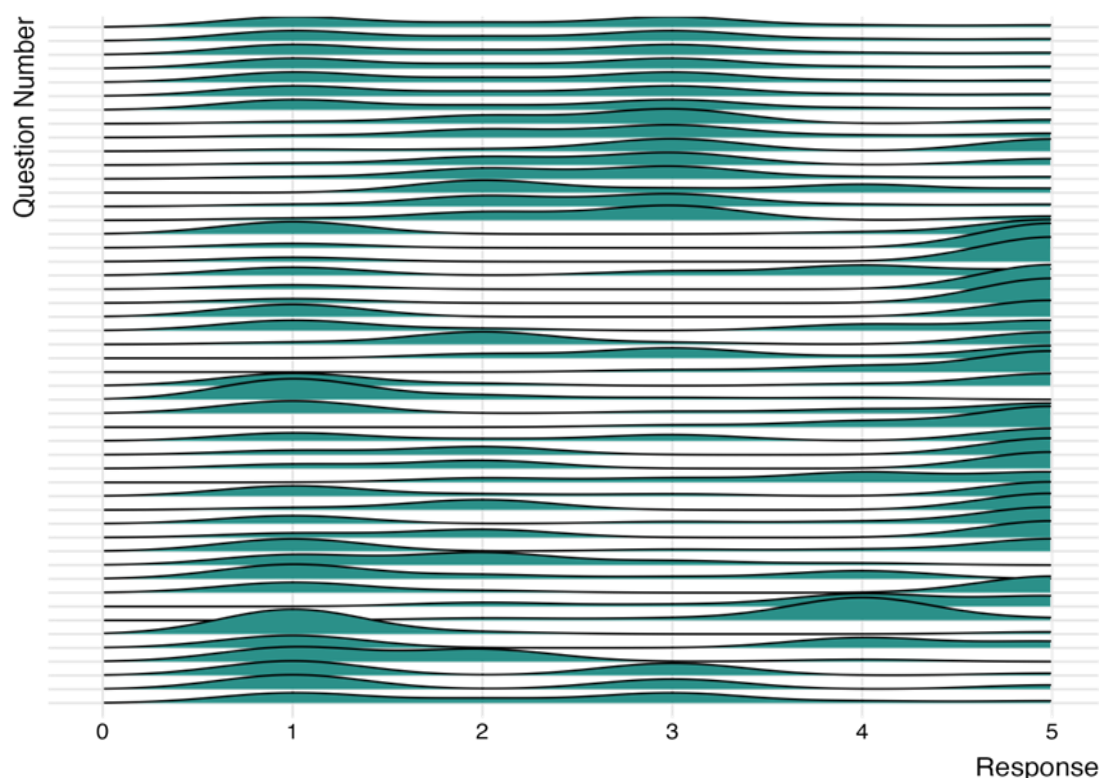


Figure 3. Ridgeline plot detailing the distribution of responses to survey questions. Questions are ordered from 1 – 50 from top to bottom (n = 258)

4.2. Correlation and Reliability Analysis

The survey comprised six constructs, each with two sub-constructs. A correlation matrix (Figure 4) revealed collinearity between responses, particularly for questions 15-18 (wellbeing measures) and 41-50 (AI strategies), indicating alignment between technology attitudes and wellbeing concerns. Kaiser-Meyer-Olkin analysis returned values of 0.5, below the threshold (<0.8) required for meaningful factor extraction.

Figure 4 reveals strong positive correlations (dark blue) between questions 15-18 (wellbeing measures) and 41-50 (AI strategies), indicating that attitudes toward technology solutions are closely aligned with teacher wellbeing concerns. The observed collinearity justifies our subsequent dimensional reduction approach through PCA.

Each of the six constructs feature two sub-constructs. Variables that inform further regression analyses were assessed using Cronbach's alpha tests for internal consistency at the sub-construct level. Using the threshold $\alpha > 0.9$ threshold for excellent internal consistency, $0.8 < \alpha < 0.9$ for good internal consistency, and $0.7 < \alpha$ for acceptable internal consistency; sub-construct work related activities during official working hours only met the threshold of acceptable internal consistency in Questions 1 and 3, with α scores of 0.75 and 0.82 respectively. For the sub-construct work related activities outside of official working hours, internal consistency was much greater with all variables bar Question 9 exceeding the threshold for good internal consistency.

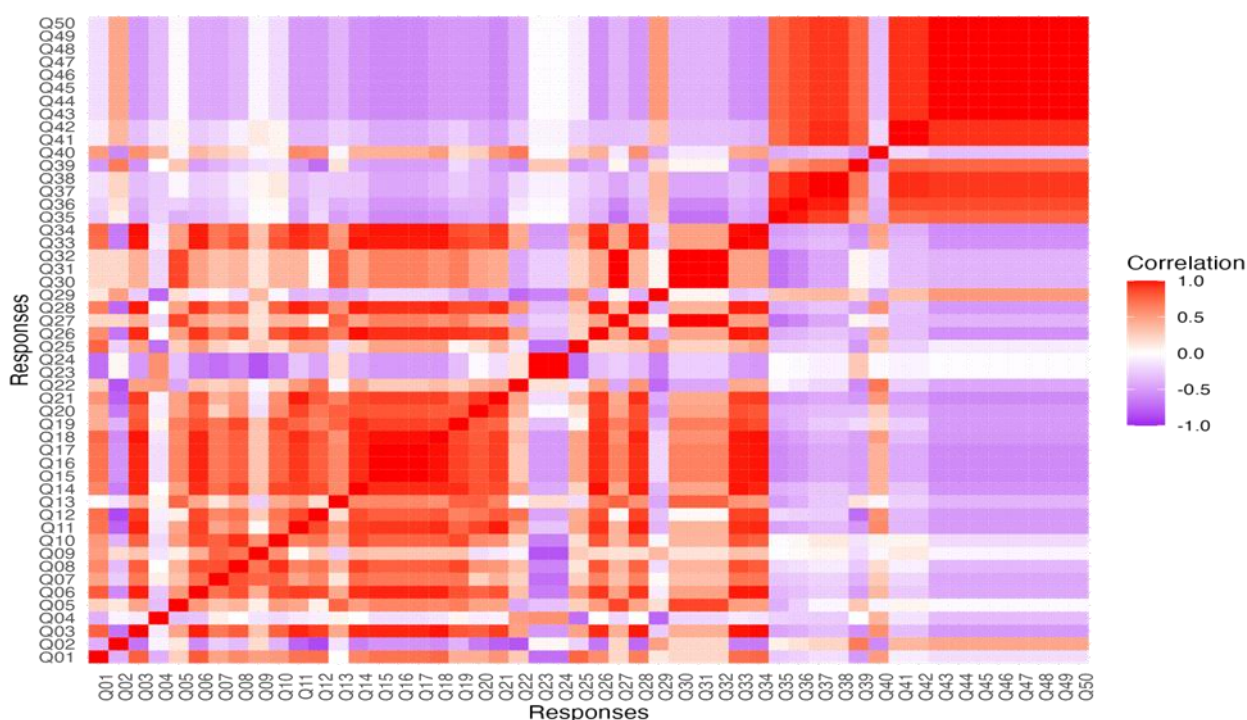


Figure 4. Correlation matrix of responses to each of the survey questions.

When considering the sub-construct sources of workload and stress, Question 13 responses are an outlier in that it is the only response to fall below the threshold of good internal consistency ($\alpha = 0.46$). In the sub-construct impact of workload on well-being, all response variables attained α scores in excess of 0.99 indicating near perfect internal consistency.

In the case of sub-construct impact of quality assurance standards, internal consistency can be judged as excellent, with only responses to Question 19 and 22 falling short of the excellent threshold at 0.70 and 0.42, respectively. The paired sub-construct teacher perspectives on balancing standards in teaching, a lower threshold of good is met for responses to Questions 23 and 24 at $\alpha = 0.72$ for both. Responses to Questions 25 and 26 were negatively correlated and therefore reversed, rendering α scores of 0.51 and 0.31 respectively.

Similar negative correlations and required reversals were discovered in the challenges in providing feedback sub-construct. Here, no response variables exceeded the threshold for acceptable internal consistency, with a highest α of 0.60. The paired sub-construct of emotional impact of feedback challenges yielded acceptable internal consistency with α scores of 0.70, 0.70, 0.80 and 0.80 for responses to Questions 31, 32, 33, and 34, respectively.

For the effectiveness of AI Tools, excellent external consistency was achieved, with α scores of 0.93, 0.99, 0.97 and 0.97 for responses to Questions 35, 36, 37 and 38, respectively. However, the sub-construct of teacher perception of AI tools yielded lesser degrees of internal consistency, with α scores of 0.85, 0.80, and 0.80 for responses to Questions 39, 41, and 42, respectively. Responses to Question 40 were negatively correlated and required reversal, after which they attained an α score of 0.33.

Responses to the remaining questions across the sub-constructs adopted strategies for feedback and impact on teacher well-being and student outcomes were both perfectly correlated with α scores of 1, indicating a perfect degree of internal consistency. Cronbach's alpha scores for each variable, assessed at the sub-construct level, are presented in the Supplementary Table (Appendix IV).

4.3. Differences Across Demographics

Kruskal-Wallis tests found no significant differences between male and female respondents except for Question 13. However, significant differences emerged across age groups, qualification levels, teaching experience, and class sizes. Figure 5 highlights that teachers with larger classes (>26 students) reported higher workload stress (Q11-14) and more negative well-being outcomes (Q15-18) compared to those with smaller classes. Additionally, early-career teachers demonstrated greater openness to AI tools (Q35-42) than their more experienced counterparts.

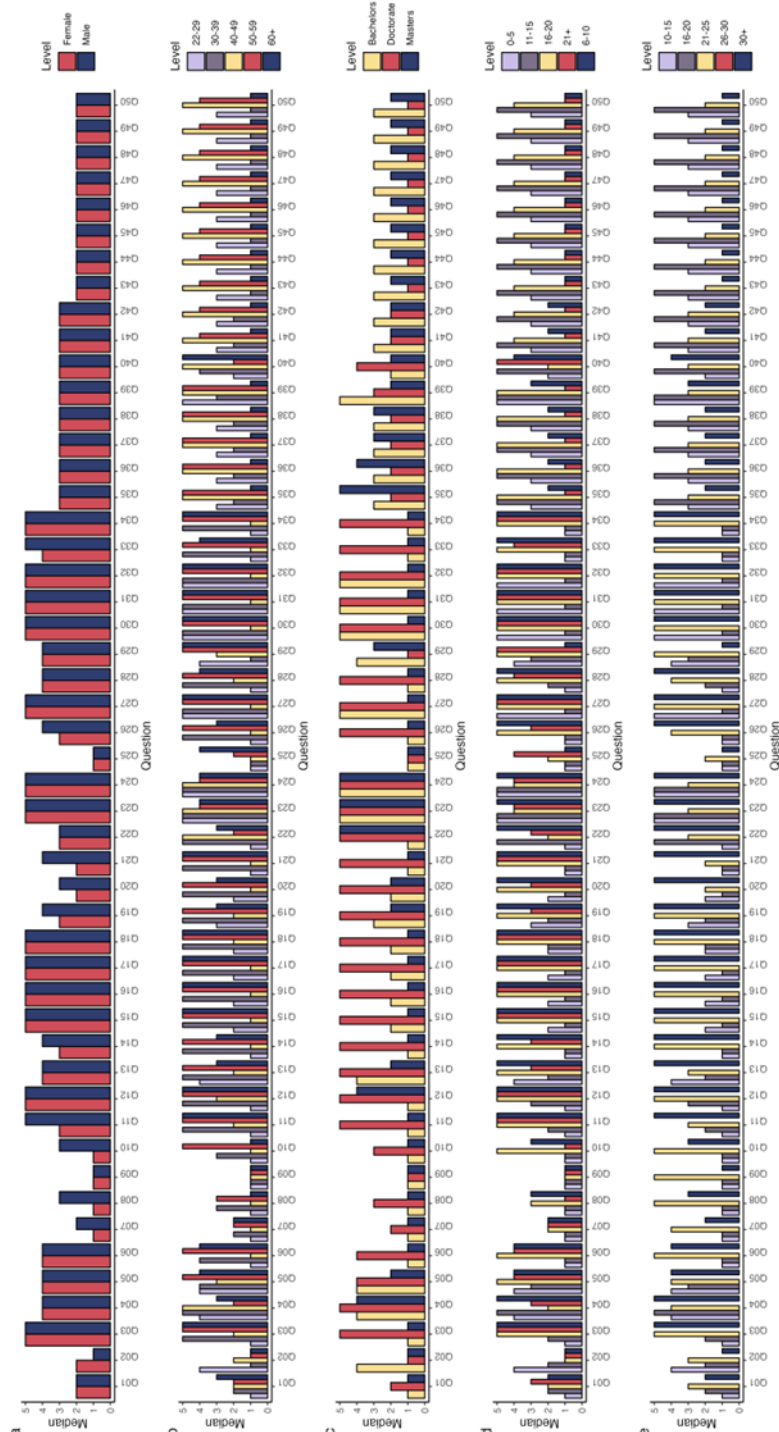


Figure 5. Median responses to Questions 1 to 50 stratified by demographic and groupings (n = 258)

4.4. PCA and LASSO Regression Analysis

LASSO regressions assessed which variables best predicted teacher well-being, job satisfaction, and ability to provide detailed feedback. LASSO applies parameterization to reduce collinearity across variables (Figure 2), choosing one variable from collinear groups while setting other coefficients to 0 to produce the simplest explanatory model. The λ value was determined using cross-validation based on the lowest RMSE value.

Initial OLS regressions with Variable Inflation Factors (VIF) revealed extensive collinearity, with most variables aliased as linear combinations of one another. We therefore employed PCA to reduce dimensionality. Scree plots (Figure 6a) showed that seven principal components were required to explain 100% of the variance, with component loadings shown in Figure 6b.

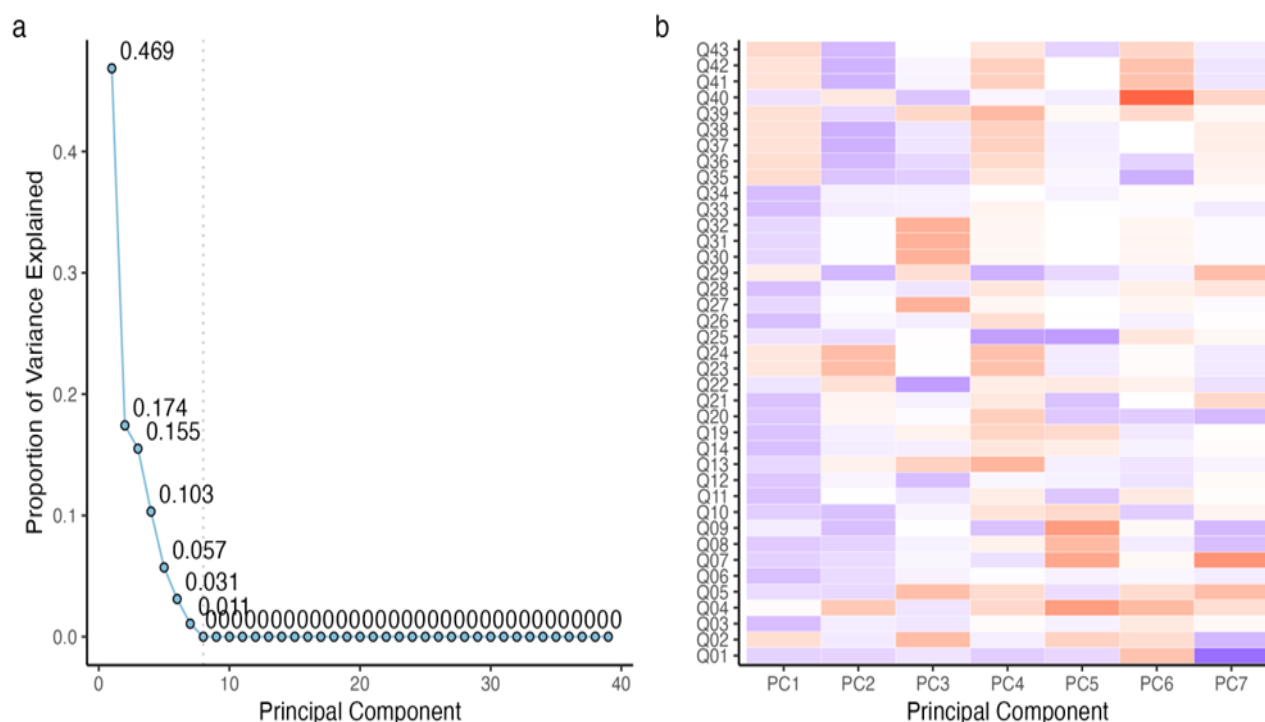


Figure 6. Wellbeing PCA parameters: a) a scree plot indicating the cumulative variance captured as the number of principal components increases, and b) PCA loadings across question response variables.

LASSO regression investigated predictors of teacher wellbeing, taken as the response to Question 16: "To what extent do you feel that your workload affects your emotional wellbeing?" and job satisfaction (Question 17). Given their high correlation ($\alpha=0.99$), one variable was removed before analysis. Cross-validation determined optimal λ values (Figure 7).

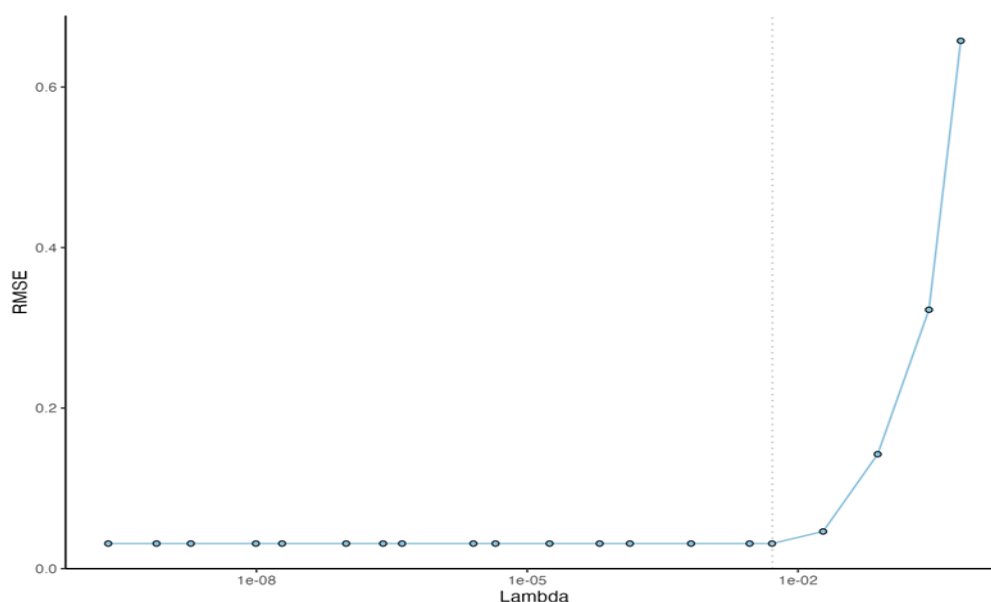


Figure 7. Lambda value selection for LASSO linear regression parameterization in search of the best predictors for teacher well-being and job satisfaction.

Outputs revealed six principal components significantly contributed to teacher wellbeing and job satisfaction (Table 3). Of these components, three were increasingly close to zero and had diminished influence. The strongest predictors of teacher wellbeing and job satisfaction are Principal Components 2, 4, and 6. PC2 has significant loadings in the work-related activities, AI-assisted tools, and innovative strategies and technologies constructs. These strong loadings are negatively correlated, indicating that low Likert scale responses in these constructs lead to greater wellbeing and job satisfaction.

To interpret these principal components more meaningfully, Principal Component 2, which strongly predicted wellbeing, primarily represents administrative workload variables (loading values >0.7 for documentation time, form completion, and reporting requirements). This suggests that administrative tasks, rather than direct teaching activities, may be particularly detrimental to teacher wellbeing. PC4 is primarily composed of variables related to teaching tasks outside official working hours (Questions 6-9), indicating that work-life boundary erosion significantly impacts wellbeing. PC6 primarily captures assessment burden variables (Questions 28-30), suggesting that the feedback intensity particular to language teaching represents a distinct stress factor.

Table 3. LASSO linear regression output for principal component predictors of teaching well-being and satisfaction.

Term	Coefficient
(Intercept)	-9.78E-18
PC1	-2.28E-01
PC2	-4.08E-02
PC3	1.87E-02
PC4	3.94E-03
PC5	-3.18E-02
PC6	1.02E-02

An identical approach was employed to investigate the most suitable predictors for the ability to provide good feedback. Cronbach alpha scores for this sub-construct were not acceptable and variables can therefore not be combined or excluded. Question 29: “How much time do you typically spend providing detailed written corrective feedback on a single student’s writing assignment?” can be used as a proxy of the ability to provide good feedback, given the quality of given feedback is likely to correlate with time spent providing it. PCA was re-run with Question 29 responses as the target variable yielding eight principal components that captured the variance of all question responses (Figure 8). Cross-validated λ values conducive to the lowest RMSE value are presented in Figure 9.

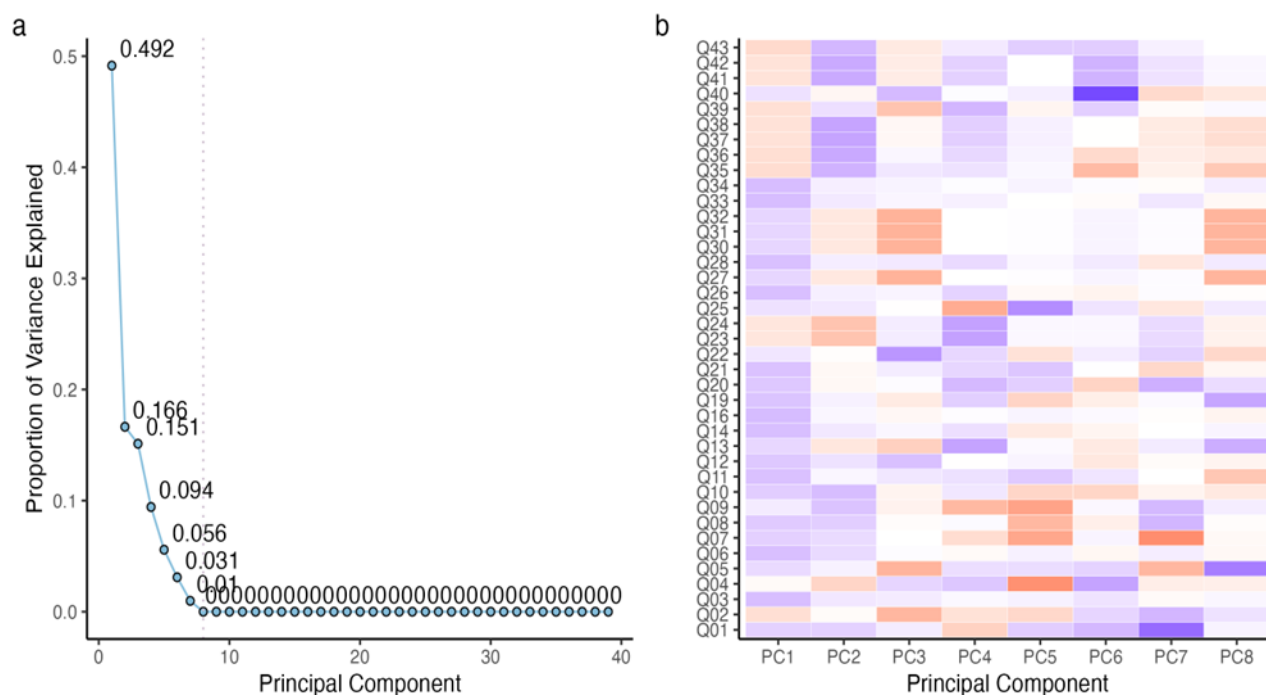


Figure 8. Ability to provide good feedback PCA parameters: a) a scree plot indicating the cumulative variance captured as the number of principal components increases, and b) PCA loadings across question response variables.

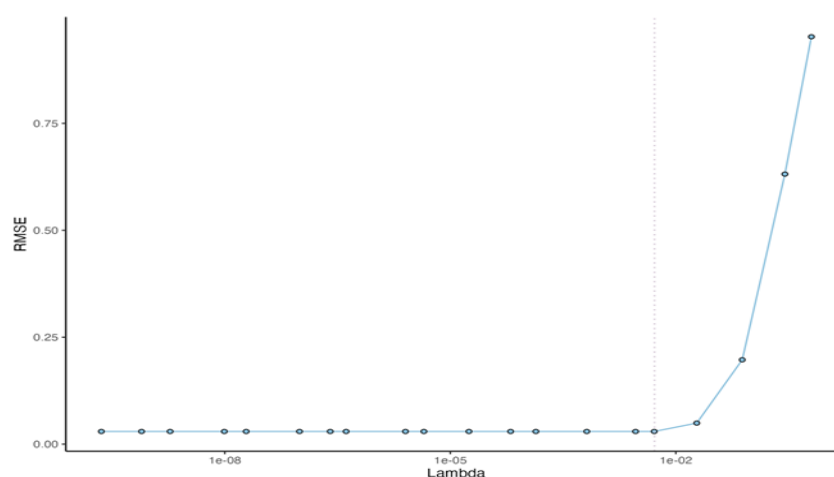


Figure 9. Lambda value selection for LASSO linear regression parameterisation in search of the best predictors for the ability to provide good feedback.

Seven of the eight principal components were shown to exert influence on the ability to provide good feedback (Table 3). PC4 and PC7 are the most influential components in the model, with the largest positive coefficients (0.322 and 0.360, respectively). PC4 is highly influenced by negative loadings of AI-assisted tools, and innovative strategies and technologies constructs, where PC7 is

influence by positive loadings in responses to Questions 5 and 7, as well a slightly positive loadings of responses within the effectiveness of AI tools sub-construct. They explain a significant amount of variation in the ability to provide good feedback and positively contribute to its prediction. PC5 is the only component with a negative coefficient (-0.176).

Interpreting these components in relation to feedback provision: PC4, with its negative loadings on AI tools, suggests that teachers skeptical of technological solutions currently spend more time on feedback, possibly compensating for perceived inadequacies of automated systems. PC7's strong association with Questions 5 and 7 (weekend and evening work) indicates that teachers providing more thorough feedback are doing so by extending work into personal time. The negative coefficient for PC5, which primarily represents variables related to quality assurance activities, confirms our SEM findings that administrative requirements directly diminish feedback capacity.

Table 3. LASSO linear regression output for predictors of providing good feedback

Term	Coefficient
(Intercept)	-1.43E-16
PC1	6.12E-02
PC2	1.72E-01
PC3	1.77E-01
PC4	3.22E-01
PC5	-1.76E-01
PC6	5.52E-02
PC7	3.36E-01

Structural equation modeling (SEM) was selected for its capacity to simultaneously test multiple relationships between observed and latent variables while accounting for measurement error. This approach allows us to examine both direct and indirect effects within a single comprehensive model, providing insight into the complex interrelationships between quality assurance standards, workload, and feedback capacity. The model specification was guided by both theoretical considerations from the Job Demands-Resources framework and preliminary correlation analyses suggesting mediation effects.

4.5. Structural Equation Modelling

Structural equation modelling was used to test the impact of quality assurance standards on workload and feedback provision. Here, the responses of Questions 20, 21 and 22, were combined as a measure of quality assurance impact - justified by the Cronbach alpha scores outlined in Section 3.1. Workload, measured by Questions 11, 12 and 14, and feedback provision - measured by Questions 27, 28, 29 and 30, were specified as latent variables within the model. The developed model posits that quality assurance directly influences both workload and feedback provision. Covariance between workload a feedback was also considered indicating that these two constructs may be related but not necessarily causally.

The Diagonally Weight Least Squares (DLWS) method was used as an estimator due to its suitability for Likert scale and ordinal data. Chi-square results of 1.274 (df = 4, p = 0.866) indicate a good model fit and no significant deviations from the data. Both the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) scored values > 0.95 at 0.998 and 0.994, respectively, indicating an excellent model fit. The Normed Fit Index (NFI) was 0.996, exceeding the 0.90 threshold for acceptable fit

(Bentler & Bonett, 1980). An RMSEA score of 0.017 further indicated excellent model fit, well below the 0.08 threshold for adequate fit (Browne & Cudeck, 1993). All latent variables had strong factor loadings in excess of 0.87.

Regressions within the SEM revealed a strong and positive effect of quality assurance standards on workload (Estimate = 0.388, $p < 0.001$) and a strong and significant negative effect of on feedback provision (Estimate = -0.224, $p < 0.001$). Further, a strong positive correlation exists between workload and feedback provision (Covariance = 0.654, $p < 0.001$). Overall, the model suggests that QA standards have a dual impact (Figure 10), increasing workload while decreasing the time available to provide good feedback. These coefficients indicate that for each unit increase in quality assurance requirements, teacher workload increases by 0.388 units while feedback capacity decreases by 0.224 units. In practical terms, this means that as institutions implement more stringent documentation requirements, teachers simultaneously experience heavier workloads and reduced ability to provide the personalized feedback essential for language acquisition. The strong covariance (0.654) between workload and feedback provision suggests these are deeply interconnected aspects of teachers' professional experience, creating a challenging dynamic where quality assurance mechanisms intended to improve educational outcomes may paradoxically undermine them."

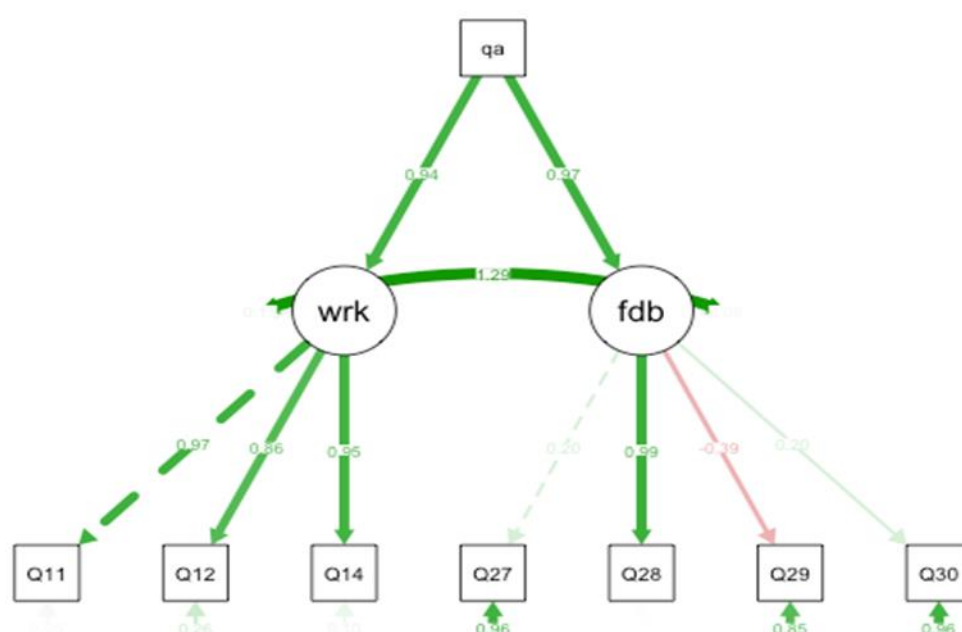


Figure 10. SEM path of the impact of quality assurance standards on teacher workload and time spent on providing written feedback.

4.6. Mediation Analysis

Mediation analysis was used to investigate the relationship between workload, quality assurance standards and wellbeing, via linear regression approaches. Workload was computed as the sum of responses to Questions 11, 12 and 14; wellbeing was derived from the sum of responses to Questions 15, 16, 17, and 18; and quality assurance was calculated as the sum of responses to Questions 20, 21 and 22 - with Cronbach's alpha scores detailed in Section 3.1 as the rationale for combination. A mediator model was parameterised by regressing workload on quality assurance standards. The outcome model was estimated to examine the impact of workload on quality assurance and wellbeing, where wellbeing is regressed against both workload and quality assurance standards to determine if the relationship is mediated.

The Average Causal Mediation Effect (ACME) represents the indirect effect of quality assurance on wellbeing through workload and suggest that higher responses to questions on QA leads to

increased workload and reduced wellbeing (Estimate = 1.859, LCL = 1.692, UCL = 2.030 $p < 0.001$). The Average Direct Effect (ADE) measures the direct effect of QA on wellbeing while controlling the effect of workload and suggests that independent of workload, high QA burden is associated with reduced teacher wellbeing (Estimate = -0.661, LCL = -0.805, UCL = -0.510, $p < 0.001$). The proportion mediated provides a ratio of the mediated effect to the total effect. An estimate greater than 1 indicates that the mediation effect is larger than the direct effect. Here, results suggest that a substantial portion of the total effect of QA on wellbeing is mediated through workload (Estimate = 1.549, LCL = 1.504, UCL = 1.710, $p < 0.001$).

These mediation results reveal important causal pathways in the relationship between quality assurance, workload, and wellbeing. The proportion mediated value exceeding 1 (1.549) indicates a suppression effect, where quality assurance's direct and indirect effects work in opposite directions. Practically speaking, while increasing quality assurance directly reduces wellbeing to some degree (-0.661), its much stronger negative effect occurs by significantly increasing workload (1.859), which then substantially reduces wellbeing. This finding suggests that institutional interventions targeting the workload implications of quality assurance frameworks would be more effective than modifying the frameworks themselves.

4.7. Current practices vs potential AI tools

Paired Wilcoxon Signed Rank tests were used to assess if there is a difference in perception between the sub-constructs of effectiveness of AI tools and teacher perceptions of AI tools. Responses to questions within sub-constructs were combined with the exception of Question 40, again on the basis of results of the Cronbach's alpha tests described in Section 3.1.

Results confirm a significant difference between the sub-constructs effectiveness of AI tools and teacher perceptions of AI tools ($V = 33,411$, $p < 0.001$). The median difference between the paired observations is 1, implying that, on average, participants rated the potential of AI tools 1 point higher than their current perceptions. This can be interpreted as teachers perceptions of how effective AI tools are, are more favourable than their openness and concerns about using them (Figure 11).

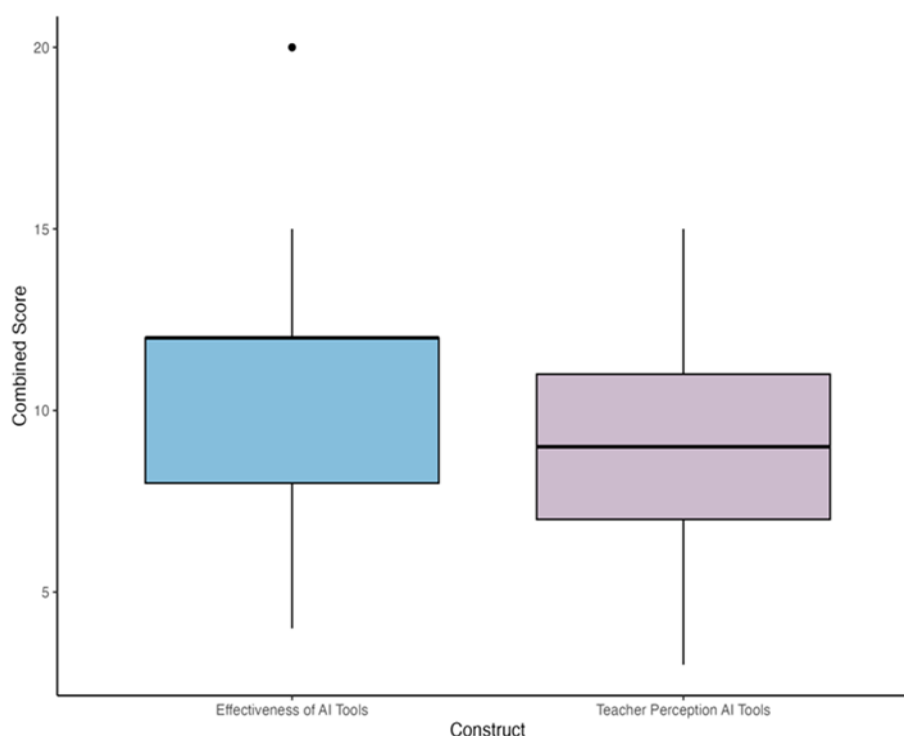


Figure 11. Median differences between effectiveness of AI tools and teacher perception of AI tools sub-constructs ($n = 258$)

4.8. Qualitative Thematic Analysis

Thematic analysis of interview and focus group data revealed five primary themes aligned with our research questions. The coding structure and theme development are presented in Table 4.

Table 4. Thematic Analysis: Codes, Categories, and Themes

Theme	Categories	Initial Codes	Representative Quote
Administrative Burden Displacement	Documentation demands; Time allocation conflicts; Pedagogical displacement	QA paperwork; Form completion; Reporting requirements; Lost teaching time; Administrative priority	"Every hour spent on documentation is an hour lost with students"
Feedback-Class Size Tension	Mathematical impossibility; Quality-quantity trade-offs; Time constraints	Large classes; Insufficient time; Feedback depth; Student numbers; Time calculations	"With 38 students... it's mathematically impossible"
Emotional Labor of Assessment	Motivation balance; Critical feedback; Psychological toll	Student demotivation; Honest critique; Emotional exhaustion; Assessment stress	"I'm constantly worried about demotivating students"
Technology-Identity Concerns	Professional role fears; Human connection; AI skepticism	Role diminishment; Personal relationships; Technology resistance; Professional identity	"I worry about the loss of personal connection"
Career-Stage Adaptation	Experience-based coping; Strategy development; Selective approaches	Feedback selectivity; Pattern focus; Efficiency strategies; Professional learning	"I've learned to be more selective with feedback"

First, administrative burden displacement emerged as teachers described how QA documentation diverted time from core teaching activities: "Every hour spent on documentation is an hour lost with students. Paperwork has become the priority rather than the pedagogy" (Participant 2, experienced). Second, feedback-class size tension highlighted the mathematical impossibility of providing detailed feedback with increasing class sizes: "With 38 students and weekly writing assignments, it's mathematically impossible to provide the detailed feedback each student deserves" (Participant 11, early-career). Third, emotional labor of assessment captured the psychological toll of balancing honest critique with student motivation: "I'm constantly worried about demotivating students while still being honest about their mistakes" (Participant 8, mid-career). Fourth, technology-identity concerns revealed how resistance to AI stemmed from fears of professional role diminishment: "I can see the potential benefits of AI, but I worry about the loss of personal connection with students" (Participant 5, doctorate holder). Fifth, career-stage adaptation showed

differential coping strategies: "Over time, I've learned to be more selective with feedback—focusing on patterns rather than every error" (Participant 9, mid-career).

Inter-rater reliability for the coding process achieved 84% agreement between independent coders. These themes complemented quantitative findings by providing explanatory mechanisms for statistical relationships, particularly the connections between QA requirements, workload intensification, and diminished wellbeing.

5. Discussion

This section integrates findings from both quantitative and qualitative analyses to address our research questions concerning teacher burnout, workload intensification, quality assurance demands, feedback challenges, and the potential of AI-assisted solutions. By triangulating statistical results with thematic interview data, we present a comprehensive understanding of the burnout phenomenon among ESL/EFL teachers. Throughout this discussion, representative quotes were selected based on their clarity in illustrating quantitative patterns and through a systematic process where multiple researchers independently identified exemplary statements for each theme. Selected quotes underwent reliability checks to ensure they accurately represented broader participant perspectives rather than outlier views. This methodical approach to integrating qualitative data strengthens the validity of our mixed-methods findings and provides contextual depth to statistical relationships identified in our analyses.

5.1. Factors Contributing to Teacher Burnout

Research Question 1: What are the primary factors contributing to burnout among ESL/EFL teachers?

Our findings identify workload intensification as the predominant driver of teacher burnout, consistent with Wang (2022), Chang (2022), and An and Tao (2024). The PCA analysis (Figure 4b) reveals that principal components with strong loadings from the work-related activities and teacher workload constructs were key predictors of diminished well-being (Table 2). This quantitative evidence is substantiated by qualitative interviews where participants consistently described overwhelming workloads:

"It's not just the teaching hours, it's everything else. The paperwork, the meetings, the quality assurance documentation—it all adds up until you're working nights and weekends just to stay afloat." (Participant 14, mid-career).

Demographic analysis shows significant variation in burnout experiences across age groups (Figure 3b), qualifications (Figure 3c), years of experience (Figure 3d), and class size (Figure 3e), though minimal gender differences were observed (Figure 3a). Notably, teachers over 50 years and those with doctorate degrees reported higher burnout levels, supporting Seo's (2023) concept of "linguistic vigilance overload" among highly qualified educators who maintain exacting standards. The demographic breakdown in Table 1 reveals that half of the participants hold doctorate degrees, making this relationship particularly relevant.

Interviews illuminated how burnout manifests differently across career stages:

"Early in my career, I thought the exhaustion was normal, just part of paying my dues. Fifteen years in, I realize this isn't sustainable—the system demands more documentation every year while classes keep getting larger." (Participant 7, experienced).

These findings align with the Job Demands-Resources model (Bakker & Demerouti, 2017), demonstrating how job demands consistently outweigh available resources for many educators.

5.2. Impact of Quality Assurance Standards

RQ2: How do quality assurance standards affect teacher workload and well-being?

Our structural equation modeling (Figure 8) provided definitive evidence that quality assurance standards significantly increase teacher workload (Estimate = 0.388, $p < 0.001$) while decreasing capacity for feedback provision (Estimate = -0.224, $p < 0.001$). This quantitative result directly addresses our hypothesis that quality assurance requirements contribute significantly to teacher burnout.

The mediation analysis further revealed that workload substantially mediates the relationship between quality assurance standards and wellbeing (Estimate = 1.549, LCL = 1.504, UCL = 1.710, $p < 0.001$), empirically supporting Kenny and Fluck's (2022) "quality-autonomy dilemma." Qualitative data enriched this understanding through detailed accounts of the documentation burden:

"The irony is that these quality assurance processes were meant to improve teaching, but they've had the opposite effect. I spend so much time documenting what I'm going to do and what I did that I have less time to actually prepare effective lessons." (Participant 3, experienced)

Participants across all demographics consistently recognized the importance of quality assurance (Figure 3a-e), creating what our qualitative analysis identified as the "accountability paradox"-teachers value quality standards but find the implementation bureaucratic and counterproductive. This perfectly illustrates Schaufeli et al.'s (2023) "quality paradox," where mechanisms designed to improve educational standards inadvertently undermine instructional effectiveness.

The estimated 8.7-14.2 hours weekly that ESL/EFL teachers spend on QA documentation (Almosa & Alzahrani, 2022) aligns with our participants' experiences, as reflected in both survey responses and interview accounts of administrative burden.

5.3. Feedback Challenges and Class Size

RQ3: What challenges do teachers face in providing feedback, and how do these contribute to burnout?

Our findings confirm the hypothesis that increasing class sizes and feedback demands contribute significantly to teacher burnout. With 49% of participants supervising classes of over 26 students (Table 1), many educators are operating beyond the 25-student threshold that Zhang and Hyland (2022) and Barrot (2023) identified as the point where individualized feedback becomes significantly compromised.

Quantitative data showed clear correlations between class size and both workload perception and diminished well-being (Figure 3e). This was elaborated in interviews where participants described the "feedback dilemma." One early-career teacher explained, "With 38 students and weekly writing assignments, it's mathematically impossible to provide the detailed feedback each student deserves. I have to choose between sleep and giving proper feedback" (Participant 11).

The emotional labor of providing feedback emerged as a significant factor in our analysis, complementing Bodenheimer and Shuster's (2019) concept of "assessment emotional labor" as a burnout contributor. Interestingly, our data revealed that teachers with 11-15 years of experience reported a less emotional impact from feedback challenges than other groups (Figure 3d), which qualitative findings suggest may be related to the development of coping strategies. As one mid-career teacher noted, "Over time, I've learned to be more selective with feedback--focusing on patterns rather than every error. It took years to realize I wasn't helping students by marking everything" (Participant 9).

Time-motion data from interviews revealed that teachers spend approximately 25-35 minutes per student paper, consistent with Rajab et al.'s (2016) finding of 20-40 minutes per paper. For the 49% of our participants with class sizes exceeding 26 students, this translates to unsustainable feedback hours.

5.4. Perceptions and Potential of AI Tools: Bridging the Adoption Gap

RQ4: How do teachers perceive AI tools, and what potential do these tools have for reducing burnout? Our analysis revealed a significant disconnect between perceptions of AI effectiveness and willingness to implement these tools (Figure 9, $V = 33,411$, $p < 0.001$), exemplifying what Guo and Wang (2023) termed the "AI Trust Paradox." This gap reflects what our qualitative data revealed as 'pedagogical identity threat'--fears that AI adoption might fundamentally alter the teacher's role. This extends beyond mere technical skepticism by highlighting how professional identity concerns create resistance, particularly among those with doctoral qualifications (Figure 3c). Our analysis identified three distinct dimensions of this resistance: (1) concerns about algorithmic reliability in capturing linguistic nuance, (2) fears of diminished professional authority, and (3) anxiety about technology-mediated distancing from the student-teacher relationship. These dimensions varied systematically across demographic categories, with early-career teachers primarily concerned with reliability while experienced educators emphasized relationship concerns.

Qualitative data illuminated these concerns:

"I can see the potential benefits of AI for handling routine feedback tasks, but I worry about the loss of personal connection with students. My feedback isn't just about correcting errors--it's about building relationships and understanding individual learning journeys." (Participant 5, doctorate holder)

This sentiment reflects concerns about what we term "pedagogical displacement"--fears that technological efficiency might come at the cost of the human elements that many teachers consider central to effective language instruction. However, our findings also revealed a potential pathway to adoption through framing AI as an assistant rather than a replacement:

"If AI could handle the initial grammar and vocabulary feedback--the routine aspects--I could focus more on higher-order concerns like argumentation and cultural appropriateness. That would be valuable, but only if I maintain the final say." (Participant 8, mid-career)

This teacher-centered framing aligns with Colonna's (2023) "teacher-in-the-loop" approach, where AI serves as a collaborative tool under the guidance of educators rather than an autonomous system. Principal components 4 and 7 (Figure 6b), which strongly predicted ability to provide quality feedback (Table 3), suggest that AI tools capable of reducing administrative workload while supporting routine feedback tasks could significantly enhance teachers' capacity for more meaningful pedagogical work.

The demographic variations in AI receptivity — particularly among early-career teachers, who show greater openness — suggest the need for differentiated approaches to technology integration. Focus group discussions revealed that experienced teachers became more receptive when presented with concrete examples of how AI could reduce quality assurance documentation burden:

"If AI could help with the QA paperwork rather than the teaching itself, that would address my biggest pain point. I spend hours documenting what I've already done instead of preparing for what I need to do next." (Participant 22, experienced)

These insights suggest that AI implementation strategies should prioritize administrative relief before moving to pedagogical applications, building trust incrementally while respecting teachers' professional autonomy.

5.5. Demographic Factors Influencing Burnout and Technology Adoption

RQ5: What demographic factors influence teacher burnout and openness to AI solutions?

Our quantitative analysis revealed significant variations in burnout experiences and technology acceptance across demographics, addressing our research question about individual factors. Age emerged as a strong predictor, with teachers over 50 reporting higher burnout levels but showing greater resistance to technological interventions (Figure 3b).

Educational qualifications also proved significant, with doctorate holders (50% of participants, as shown in Table 1) reporting more challenges with feedback provision (Figure 3c) while simultaneously expressing greater skepticism toward AI solutions. These findings support the hypothesis that demographic factors significantly influence both burnout experiences and attitudes toward potential solutions.

Teaching experience revealed an intriguing pattern: burnout generally increased with years in the profession (Figure 3d), except for the distinctive mid-career resilience regarding emotional impacts of feedback challenges. This pattern was elaborated in interviews:

"There's a sweet spot around year 10 where you've developed efficient systems but haven't yet been ground down by the cumulative stress. After 20 years, the constant changes in requirements become harder to adapt to." (Participant 18, experienced).

The inverse relationship between educational qualification and technology acceptance requires particular attention, as it suggests that those with the strongest pedagogical knowledge—who might benefit most from administrative relief—also express the greatest concerns about AI integration. This pattern aligns with Self-Determination Theory's emphasis on autonomy as a core psychological need (Ryan & Deci, 2023), as doctorate holders consistently described fears of diminished professional agency in their interview responses. However, their significantly higher responses to workload challenges suggest a "technology-autonomy paradox"—where resistance to potentially helpful tools may inadvertently perpetuate the very conditions causing distress. This finding has important implications for technology implementation strategies, suggesting that emphasizing enhancement rather than replacement of professional judgment may be crucial for engaging highly qualified educators. Class size showed the most predictable relationship with burnout (Figure 3e), with larger classes consistently associated with higher workload perceptions and reduced well-being, supporting our hypothesis about class size as a burnout factor.

5.6. Theoretical and Practical Implications

Our integrated findings provide substantial empirical support for established theoretical frameworks, while also revealing novel insights into ESL/EFL teacher burnout. This section contextualizes our key findings within the broader academic landscape established in our literature review.

5.6.1. Empirical Support for the Job Demands-Resources Model

Our findings strongly validate Bakker and Demerouti's (2017) Job Demands-Resources model in ESL/EFL contexts. The identification of workload intensification as the predominant driver of burnout aligns with Wang (2022), Chang (2022), and An and Tao (2024), who established workload as a primary factor across diverse educational systems. However, our research extends this understanding by demonstrating that administrative demands have a disproportionate negative effect compared to direct teaching activities—a distinction not previously explored in ESL/EFL literature.

The strong relationship we identified between quality assurance standards and increased workload empirically validates what Schaufeli et al. (2023) conceptualized as the "quality paradox," where mechanisms designed to improve educational standards inadvertently undermine effectiveness. Our structural equation modeling provides quantitative evidence for this paradox, demonstrating how finite teacher resources must be allocated between competing administrative and pedagogical demands, resulting in what we term "resource competition."

5.6.2. Contextualizing Quality Assurance Impact Within Existing Literature

Our documentation of quality assurance frameworks simultaneously increases workload while decreasing feedback capacity, directly supporting Kenny and Fluck's (2022) "quality-autonomy dilemma." The mediation analysis, revealing that workload substantially mediates the relationship between quality assurance and wellbeing, provides empirical evidence for their theoretical proposition that accountability measures erode teacher autonomy and job satisfaction.

The finding that ESL/EFL teachers spend a substantial amount of time on QA documentation, consuming significant portions of their working hours, confirms Almosa & Alzahrani's (2022) observations about "documentation density" in language teaching contexts. Our participants' experiences of administrative burden displacement align with Ma and Liu's (2024) institutional ethnography, which documents "bureaucratic displacement"—where QA requirements divert teacher attention from pedagogical activities to administrative compliance.

5.6.3. Conservation of Resources Theory and Burnout Mechanisms

The dynamic between quality assurance requirements, workload intensification, and diminished well-being exemplifies Hobfoll et al.'s (2018) Conservation of Resources theory. Our findings demonstrate how resource depletion triggers what participants described as a "quality-wellbeing tradeoff," supporting the theory's prediction of "resource loss spirals" through compromised teaching effectiveness and reduced student engagement.

The emotional labor of providing feedback that emerged in our analysis directly supports Bodenheimer and Shuster's (2019) concept of "assessment emotional labor" as a burnout contributor. Our participants' experiences of balancing critical feedback with student motivation while managing large classes validate their findings that teachers suppress their reactions while crafting balanced responses.

5.6.4. Technology Acceptance and Professional Identity

Our identification of a substantial perception-adoption gap regarding AI tools provides empirical evidence for Guo and Wang's (2023) "AI Trust Paradox." The finding that teachers recognize AI effectiveness while resisting implementation confirms their observation that professionals acknowledge time-saving potential while fearing professional devaluation.

However, our research extends Davis's (1989) Technology Acceptance Model by identifying professional identity concerns as a significant mediating factor between perceived usefulness and adoption intention, particularly among highly qualified educators. This extends beyond Guo and Wang's (2023) findings by revealing specific demographic patterns in resistance, with early-career teachers primarily concerned with algorithmic reliability, while experienced educators emphasize the preservation of relationships.

Receptivity to AI, when framed as an assistant rather than a replacement, validates Colonna's (2023) "teacher-in-the-loop" approach. Our finding that experienced teachers became more receptive when presented with administrative relief applications supports their demonstration that this approach yields substantial workload benefits while maintaining teacher satisfaction and professional agency.

5.6.5. Comparative Analysis with Existing ESL/EFL Research

While previous ESL/EFL burnout research documented general factors (Talbot & Mercer, 2018; Zarrinabadi et al., 2023), our study provides the first empirical demonstration of how quality assurance frameworks specifically impact language teachers. Our findings contrast with Seo's (2023) focus on "linguistic vigilance overload" by revealing that administrative rather than linguistic demands constitute the primary burnout driver.

The feedback challenges we documented align with Rajab et al.'s (2016) time estimates for corrective feedback while extending their findings by demonstrating how increasing class sizes create mathematically impossible feedback demands. Our identification of the 25-student threshold where feedback quality deteriorates confirms Zhang and Hyland's (2022) and Barrot's (2023) observations while providing qualitative insights into teachers' coping strategies.

5.6.6. Self-Determination Theory and Professional Autonomy

The inverse relationship between educational qualification and technology acceptance aligns with Ryan and Deci's (2023) Self-Determination Theory emphasis on autonomy as a core psychological need. Our finding that doctorate holders express the greatest concerns about AI integration while reporting the highest workload challenges reveals what we term a "technology-autonomy paradox"—where resistance to potentially helpful tools may perpetuate the conditions causing distress.

This pattern supports Sato et al.'s (2022) finding that technologies enhancing professional autonomy are negatively correlated with burnout symptoms, suggesting that the implementation approach significantly affects well-being, as noted by Pishghadam et al. (2023).

5.6.7. Implications for Sustainable ESL/EFL Education

Our findings address the "pedagogical sustainability crisis" described by Schaufeli et al. (2023), where teachers' professional commitment conflicts with practical limitations. The identification of "motivation-practice incongruence" in our qualitative data supports their assertion that this incongruence represents a more potent predictor of burnout than workload alone.

The demographic variations in AI receptivity, particularly the greater openness of early-career teachers compared to experienced educators, suggest generational shifts in technology acceptance that align with broader patterns observed in the educational technology adoption literature within our review.

5.6.8. Practical Intervention Framework

Based on our theoretical integration and empirical findings, we propose a three-tier intervention approach grounded in the literature: (1) immediate relief through AI-assisted administrative task automation, targeting the pathway we identified between quality assurance and wellbeing; (2) medium-term recalibration of quality assurance frameworks to reduce what Kovalkova & Malkova (2021) term "documentation density" while maintaining educational standards; and (3) long-term professional development focusing on technology integration that preserves the professional autonomy emphasized by Ryan and Deci (2023).

This framework addresses both the structural issues identified by Kenny and Fluck (2022) and the technological opportunities demonstrated by Alsalem (2024), Praphan and Praphan (2023), and Zhou and Hou (2025), while respecting the professional identity concerns highlighted by our participants and supported by existing literature.

6. Conclusion

This study examines the relationship between quality assurance requirements, workload intensification, and teacher burnout among EFL/ESL instructors in Saudi higher education, revealing a "quality paradox" where QA frameworks designed to enhance standards inadvertently compromise them. Statistical modeling demonstrates a significant relationship between QA requirements and diminished teacher well-being, primarily mediated through workload intensification. Qualitative findings reveal what participants described as an "accountability paradox"—valuing quality standards while finding their implementation bureaucratic and counterproductive.

Our contributions include: (1) empirical demonstration of how QA frameworks increase workload while reducing capacity for personalized feedback; (2) identification of a significant perception-adoption gap regarding AI tools; (3) demographic patterns in burnout and technology acceptance; and (4) a "teacher-in-the-loop" framework preserving professional judgment. Thematic analysis revealed "pedagogical identity threat" as a key barrier to technology adoption, particularly among highly qualified teachers concerned about maintaining meaningful student connections.

Rather than viewing burnout as an individual deficiency, institutions should recognize it as a systemic outcome of conflicting demands on teacher time. For technology integration, qualitative findings support a teacher-centered approach, where AI handles routine tasks. Implementation strategies prioritize administrative relief before pedagogical applications, building trust incrementally while respecting professional autonomy.

6.1. Limitations

Methodological limitations include our cross-sectional design, which limits causal claims, potential generalizability constraints due to the Saudi context, self-selection bias, limitations of virtual data collection, and the evolving nature of AI technologies. Despite these limitations, robust sample size, methodological triangulation, and consistent patterns support the validity of our findings.

7. Suggestion

Future research should explore the longitudinal effects of AI integration, investigate the implementation gap between perceived usefulness and actual adoption, and examine how different institutional contexts influence the effectiveness of interventions.

We recommend: (1) recalibrating QA frameworks while implementing AI solutions in three phases; (2) differentiating technology integration by career stage; (3) establishing acceptable class size thresholds; and (4) recognizing that technological solutions complement but cannot replace sustainable teaching conditions.

As ESL/EFL education faces increasing class sizes and administrative demands—trends unlikely to reverse as institutions balance cost pressures with QA requirements—supporting teacher wellbeing becomes crucial. This situation is particularly acute in non-English speaking countries pursuing international accreditations, where English teachers often bear additional QA responsibilities due to their language proficiency. The "teacher-in-the-loop" approach offers a pathway that balances technological efficiency with professional autonomy, while preserving the human connections essential to language acquisition.

Declarations

Author Contributions. (Sole Author) Conceptualization, methodology, data collection, formal analysis, writing—original draft preparation, writing—review and editing.

Conflicts of Interest. The author declares no conflict of interest.

Funding. This research received no external funding.

Ethical Approval. This study was conducted in accordance with established ethical principles for educational research and was approved by the Ethics Committee of the institutions where the study was conducted. Informed consent was obtained from all participants involved in the study. All data were anonymized to protect participant privacy.

Data Availability Statement. The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy considerations for the participants. The data collection instruments are all appended to this manuscript as appendices.

Acknowledgments. The authors would like to thank the ESL/EFL teachers who participated in this study for their time and valuable insights. The authors also extend their gratitude to the editors and anonymous peer reviewers for their constructive feedback, which greatly enhanced the clarity and quality of this manuscript.

References

- Agyapong, B., Obuobi-Donkor, G., Burbach, L., & Wei, Y. (2023). Burnout and Associated Psychological Issues Among Teachers: A Scoping Review. *European Psychiatry*, 66, S948 - S949. <https://doi.org/10.1192/j.eurpsy.2023.2010>
- Alghamdy, R. Z. (2023). Pedagogical and ethical implications of artificial intelligence in EFL context: A review study. *English Language Teaching*, 16(10), 87-98. <https://doi.org/10.5539/elt.v16n10p87>
- Almossa, S. Y., & Alzahrani, S. M. (2022). Transnational English Language Teachers' Assessment Practices in Higher Education TESOL Landscape. In *Transnational English Language Assessment Practices in the Age of Metrics* (pp. 93-106). Routledge. <https://doi.org/10.4324/9781003252382-9>
- Alsalem, M. S. (2024). EFL teachers' perceptions of the use of an AI grading tool (CoGrader) in English writing assessment at Saudi universities: an Activity Theory Perspective. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2430865>
- Alshahrani, A., & Storch, N. (2025). Investigating the effectiveness of scaffolded feedback on EFL Saudi students' writing accuracy: A longitudinal classroom-based study. *Assessing Writing*, 63, 100910. <https://doi.org/10.1016/j.asw.2024.100910>
- Alutaybi, M. M., & Alfares, N. S. (2024). EFL teachers' perspective of teaching large online classes: Issues and challenges. *Journal of Language Teaching and Research*, 15(1), 255-262. <https://doi.org/10.17507/jltr.1501.28>
- An, S., & Tao, S. (2024). English as a foreign language teachers' burnout: The predictor powers of self-efficacy and well-being. *Acta psychologica*, 245, 104226. <https://doi.org/10.1016/j.actpsy.2024.104226>
- Bakker A. B., Demerouti E. (2017). Job demands-resources theory: Taking stock and looking forward. *Journal of Occupational Health Psychology*, 22(3), 273-285. <https://doi.org/10.1037/ocp0000056>
- Barrot, J. S. (2023). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, 36(4), 584-607. <https://doi.org/10.1080/09588221.2021.1936071>
- Bodenheimer, G., & Shuster, S. M. (2019). Emotional labour, teaching and burnout: Investigating complex relationships. *Educational Research*, 62(1), 63-76. <https://doi.org/10.1080/00131881.2019.1705868>

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Chang, H. (2022). Stress and burnout in EFL teachers: The mediator role of self-efficacy. *Frontiers in psychology*, 13, 880281. <https://doi.org/10.3389/fpsyg.2022.880281>.
- Chen, F., Wang, X., & Gao, Y. (2024). EFL teachers' burnout in technology-enhanced instruction settings: The role of personality traits and psychological capital. *Acta Psychologica*, 249, 104461. <https://doi.org/10.1016/j.actpsy.2024.104461>.
- Chen, J., Lin, C., & Lin, F. (2024). The Interplay among EFL Teachers' Emotional Intelligence, Self-Efficacy, and Burnout *Acta Psychologica*, 248, 104364. <https://doi.org/10.1016/j.actpsy.2024.104364>.
- Chen, X., Wang, X., & Qu, Y. (2023). Constructing ethical AI based on the “Human-in-the-Loop” system. *Systems*, 11(11), Article 548. <https://doi.org/10.3390/systems11110548>
- Colonna, L. (2023). Teachers in the loop? An analysis of automatic assessment systems under Article 22 GDPR. *International Data Privacy Law*, 14(1), 3–18. <https://doi.org/10.1093/idpl/ipad024>
- Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage Publications.
- Dakakni, D., & Safa, N. (2023). Artificial intelligence in the L2 classroom: Implications and challenges on ethics and equity in higher education: A 21st century Pandora’s box. *Computers and Education: Artificial Intelligence*, 5, 100179. <https://doi.org/10.1016/j.caeai.2023.100179>
- Davis, F. (1989). Technology Acceptance Model: Origins. *Working Papers on Information Systems*, 35-59. <https://doi.org/10.4018/978-1-4666-8156-9.ch013>
- Derakhshan, A., Eslami, Z. R., Curle, S., & Zhaleh, K. (2022). Exploring the validity of immediacy and burnout scales in an EFL context: The predictive role of teacher-student interpersonal variables in university students’ experience of academic burnout. *Studies in Second Language Learning and Teaching*, 12(1), 87–115. <https://doi.org/10.14746/ssl.t.2022.12.1.5>
- Ding, L. (2024). Exploring the causes, consequences, and solutions of Chinese EFL teachers' psychological ill-being: A qualitative investigation. *European Journal of Education*, 59(4), e12739. <https://doi.org/10.1111/ejed.12739>.
- Droogenbroeck, F., Spruyt, B., & Vanroelen, C. (2014). Burnout among senior teachers: Investigating the role of workload and interpersonal relationships at work. *Teaching and Teacher Education*, 43, 99-109. <https://doi.org/10.1016/J.TATE.2014.07.005>.
- Ferris, D. R. (2022). Feedback on L2 student writing: Current trends and future directions. In *Handbook of practical second language teaching and learning* (pp. 344-356). Routledge. <https://doi.org/10.4324/9781003106609-28>
- Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health Services Research*, 48(6pt2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage Publications.
- Flom, P. L., & Cassell, D. L. (2007). Stopping stepwise: Why stepwise and similar selection methods are harmful, and what you should use. In NorthEast SAS Users Group Inc 20th Annual Conference: 11-14th November 2007 (pp. 1-6).

- Guo, K., & Wang, D. (2024). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 29(7), 8435-8463. <https://doi.org/10.1007/s10639-023-12146-0>
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.
- Hobfoll S. E., Halbesleben J., Neveu J.-P., Westman M. (2018). Conservation of resources in the organizational context: The reality of resources and their consequences. *Annual Review of Organizational Psychology and Organizational Behavior*, 5, 103–128. <https://doi.org/10.1146/annurev-orgpsych-032117-104640>
- Hollander, M., Wolfe, D. A., & Chicken, E. (2014). Nonparametric statistical methods (3rd ed.). John Wiley & Sons.
- Hyland, K., & Hyland, F. (Eds.). (2019). Feedback in second language writing: Contexts and issues Cambridge, UK: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139524742>.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning with Applications in R (2nd ed.). Springer.
- Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, 38(12), 1217-1218. <https://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Jonäll, K. (2024). *Artificial intelligence in academic grading: A mixed-methods study* [Master's Dissertation, University of Gothenburg]. Gothenburg University Publications Electronic Archive. Available at: https://gupea.ub.gu.se/bitstream/handle/2077/83561/Jonäll_HPA202_VT24.pdf
- Kenny, J., & Fluck, A. E. (2022). Emerging principles for the allocation of academic work in universities. *Higher Education*, 83(6), 1371–1388. <https://doi.org/10.1007/s10734-021-00747-y>
- Ko, C. J. (2022). Online individualized corrective feedback on EFL learners' grammatical error correction. *Computer Assisted Language Learning*, 1–29. <https://doi.org/10.1080/09588221.2022.2118783>
- Kovalkova, T., & Malkova, T. (2021). Burnout Syndrome: A Study among Lecturers. Proceedings of the International Conference on Economics, Law and Education Research (ELER 2021), Vol 170, 69-74. <https://doi.org/10.2991/aebmr.k.210320.013>
- Lee, Y. (2017). Emotional labor, teacher burnout, and turnover intention in high-school physical education teaching. *European Physical Education Review*, 25, 236 - 253. <https://doi.org/10.1177/1356336X17719559>.
- Li, S., & Vuono, A. (2019). Twenty-five years of research on oral and written corrective feedback in System. *System*, 84, 93–109. <https://doi.org/10.1016/j.system.2019.05.006>
- Ma, Y., & Liu, Z. (2024). Emotion regulation and well-being as factors contributing to lessening burnout among Chinese EFL teachers. *Acta Psychologica*, 245, 104219. <https://doi.org/10.1016/j.actpsy.2024.104219>.
- Malik, A., Khan, M. L., Hussain, K., Qadir, J., & Tarhini, A. (2025). AI in higher education: unveiling academicians' perspectives on teaching, research, and ethics in the age of ChatGPT. *Interactive Learning Environments*, 33(3), 2390-2406. <https://doi.org/10.1080/10494820.2024.2409407>
- Manuel, J., Carter, D., & Dutton, J. (2018). 'As much as I love being in the classroom': Understanding secondary English teachers' workload. *English in Australia*, 53(3), 5 <https://search.informit.org/doi/10.3316/ielapa.232694249888988>

- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior*, 2(2), 99–113. <https://doi.org/10.1002/job.4030020205>.
- Naidoo, S. (2022). Adapting the technology acceptance model to investigate student responses to e-learning. *South African Journal of African Languages*, 42(1), 1-8. <https://doi.org/10.1080/02572117.2021.2015111>
- Nayernia, A., & Babayan, Z. (2019). EFL teacher burnout and self-assessed language proficiency: exploring possible relationships. *Language Testing in Asia*, 9(1), 3. <https://doi.org/10.1186/s40468-019-0079-6>.
- Pennington, M. C., & Richards, J. C. (2016). Teacher Identity in Language Teaching: Integrating Personal, Contextual, and Professional Factors. *RELC Journal*, 47(1), 5-23. <https://doi.org/10.1177/0033688216631219>
- Persico, D., Manca, S., & Pozzi, F. (2014). Adapting the technology acceptance model to evaluate the innovative potential of e-learning systems. *Computers in Human Behavior*, 30, 614-622. <https://doi.org/10.1016/j.chb.2013.07.045>
- Pike, M., Towey, D., & Walker, J. (2019). Identifying and Alleviating Assessment Stress in Higher Education. In K. C. Li, & E. Tsang (Eds.), *2019 International Conference on Open and Innovative Education (ICOIE 2019)* (pp. 589-595). The Open University of Hong Kong. <https://research.nottingham.edu.cn/files/449287278/ICOIE.2019.StressRelief.Preprint.pdf>
- Pishghadam, R., Ebrahimi, S., Golzar, J., & Miri, M. A. (2023). Introducing emo-educational divorce and examining its relationship with teaching burnout, teaching motivation, and teacher success. *Current Psychology*, 42(33), 29198-29214. <https://doi.org/10.1007/s12144-022-04000-2>
- Prapahan, P. W., & Prapahan, K. (2023). AI technologies in the ESL/EFL writing classroom: The villain or the champion?. *Journal of Second Language Writing*, 62, 101072. <https://doi.org/10.1016/j.jslw.2023.101072>.
- Pressley, T. (2021a). Factors Contributing to Teacher Burnout During COVID-19. *Educational Researcher*, 50(5), 325-327. <https://doi.org/10.3102/0013189X211004138>
- Pressley, T. (2021b). Returning to teaching during COVID-19: An empirical study on elementary teachers' self-efficacy. *Psychology in the Schools*, 58(8), 1611-1623. <https://doi.org/10.1002/pits.22528>
- Pressley, T., & Ha, C. (2022). Teacher exhaustion during COVID-19: Exploring the role of administrators, self-efficacy, and anxiety. *The Teacher Educator Journal*, 57(1), 61-78. <https://doi.org/10.1080/08878730.2021.1995094>.
- R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Rajab, H., Khan, K., & Elyas, T. (2016). A case study of EFL teachers' perceptions and practices in written corrective feedback. *International Journal of Applied Linguistics and English Literature*, 5(1), 119-131. <https://doi.org/10.7575/aiac.ijalel.v.5n.1p.119>
- Ryan, R.M., Deci, E.L. (2023). Self-Determination Theory. In: Maggino, F. (eds) *Encyclopedia of Quality of Life and Well-Being Research*. Springer, Cham. https://doi.org/10.1007/978-3-031-17299-1_2630

- Sadoughi, M., Hejazi, S., & Khajavy, G. (2024). Protecting language teachers from burnout: The roles of teaching mindset, teaching grit, and emotion regulation. *Language Teaching Research*. <https://doi.org/10.1177/13621688241238350>.
- Sato, M., Castillo, F. F., & Oyanedel, J. C. (2022). Teacher motivation and burnout of English-as-a-foreign-language teachers: Do demotivators really demotivate them? *Frontiers in Psychology*, 13, 1–12. <https://doi.org/10.3389/fpsyg.2022.891452>
- Schaufeli W. B., Hakanen J. J., Shimazu A. (2023). Burning questions in burnout research. In De Cuyper N., Selenko E., Euwema M., Schaufeli W. (Eds.), *Job insecurity, precarious employment and burnout* (pp. 127–148). Edward Elgar Publishing eBooks. <https://doi.org/10.4337/9781035315888.00017>
- Seo, Y. (2023). Enthusiasm, obsession, or delusion? language ideologies and negotiating identities of one non-native English teacher in an EFL context. *International Journal of Multilingualism*, 21(3), 1621–1634. <https://doi.org/10.1080/14790718.2023.2200253>
- Shadiev, R., & Feng, Y. (2023). Using automated corrective feedback tools in language learning: A review study. *Interactive Learning Environments*, 1–29. <https://doi.org/10.1080/10494820.2022.2153145>
- Shi H, Aryadoust V. A systematic review of AI-based automated written feedback research. *ReCALL*. 2024;36(2):187-209. <https://doi.org/10.1017/S0958344023000265>
- Siyum, B. A. (2023). University instructors' burnout: antecedents and consequences. *Journal of Applied Research in Higher Education*, 15(4), 1056-1068. <https://doi.org/10.1108/jarhe-04-2022-0131>
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91, 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Sun, X., & Zhang, W. (2023). Teacher emotions and professional development: A case study of three first-year EFL teachers. *Issues in Educational Research*, 33(3), 1190–1209. <https://search.informit.org/doi/10.3316/informit.T2024050800009590936811819>
- Tabachnick, B. G., & Fidell, L. S. (2019). Using multivariate statistics (7th ed.). Pearson.
- Talbot, K., & Mercer, S. (2018). Exploring University ESL/EFL Teachers' Emotional Well-Being and Emotional Regulation in the United States, Japan and Austria. *Chinese Journal of Applied Linguistics*, 41, 410 - 432. <https://doi.org/10.1515/CJAL-2018-0031>.
- Tashakkori, A., & Teddlie, C. (Eds.). (2010). Sage handbook of mixed methods in social & behavioral research (2nd ed.). Sage Publications.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Walker, M. E., Mehmood, R., Koshinsky, J., Hedlin, P., Chakravarti, A., Goncin, U., ... & O'Brien, J. M. (2024). Professional fulfillment, burnout, and wellness: a repeated cross-sectional survey in the COVID-19 pandemic era. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*, 71(12), 1775-1777. <https://doi.org/10.1007/s12630-024-02807-0>

- Wang, Z. (2022). The effect of EFL teacher apprehension and teacher burnout on learners' academic achievement. *Frontiers in Psychology*, 12, 839452. <https://doi.org/10.3389/fpsyg.2021.839452>.
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications. <https://doi.org/10.4135/9781412983655>
- Yin, H., Huang, S., & Chen, G. (2019). The relationships between teachers' emotional labor and their burnout and satisfaction: A meta-analytic review. *Educational Research Review*. <https://doi.org/10.1016/J.EDUREV.2019.100283>.
- Yu, S., Zheng, Y., Jiang, L., Liu, C., & Xu, Y. (2021). "I even feel annoyed and angry": Teacher emotional experiences in giving feedback on student writing. *Assessing Writing*, 48, 100528. <https://doi.org/10.1016/j.asw.2021.100528>
- Zarrinabadi, N., Jamalvandi, B., & Rezazadeh, M. (2023). Investigating fixed and growth teaching mindsets and self-efficacy as predictors of language teachers' burnout and professional identity. *Language Teaching Research*, 0(0). <https://doi.org/10.1177/13621688231151787>
- Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102. <https://doi.org/10.1016/j.asw.2018.02.004>
- Zhang, Z. V., & Hyland, K. (2022). Fostering student engagement with feedback: An integrated approach. *Assessing Writing*, 51, 100586. <https://doi.org/10.1016/j.asw.2021.100586>.
- Zhou, C., & Hou, F. (2025). How Do EFL Teachers Utilize AI Tools in Their Language Teaching?. *Theory & Practice in Language Studies (TPLS)*, 15(2). <https://doi.org/10.17507/tpls.1502.10>.
- Zhou, J., Ke, P., Qiu, X., Huang, M., & Zhang, J. (2023). ChatGPT: Potential, prospects, and limitations. *Frontiers of Information Technology & Electronic Engineering*, 25(1), 6–11. <https://doi.org/10.1631/FITEE.2300089>
- Zhou, L., Xue, S., & Li, R. (2022). Extending the Technology Acceptance Model to explore students' intention to use an online education platform at a University in China. *Sage Open*, 12(1), 21582440221085259. <https://doi.org/10.1177/21582440221085259>

About the Contributor(s)

Wael Alharbi is an Associate Professor of Applied Linguistics and Computer-Assisted Language Learning at the Yanbu English Language and Preparatory Year Institute, The Royal Commission for Jubail and Yanbu, Saudi Arabia. His research focuses on technology-enhanced language learning, digital pedagogy, and EFL education. He has published extensively in Scopus-indexed journals and actively contributes to CALL research.

Email: harbiw@rcjy.edu.sa

ORCID: <https://orcid.org/0000-0001-7713-5294>

Publisher's Note: The opinions, statements, and data presented in all publications are solely those of the individual author(s) and contributors and do not reflect the views of Universitepark, EDUPIJ, and/or the editor(s). Universitepark, the Journal, and/or the editor(s) accept no responsibility for any harm or damage to persons or property arising from the use of ideas, methods, instructions, or products mentioned in the content.

Appendix I. The Questionnaire

Background Information		1	Gender	Male [1]	Female [2]					
		2	Age group	22-29 years [1]	30-39 years [2]	40-49 years [3]	50-60 years [4]	over 60 years [5]		
		3	What is the highest degree you have obtained?	Bachelor's Degree (BA, BS, etc.) [1]	Master's Degree (MA, MS, MBA, etc.) [2]	Doctorate (Ph.D., Ed.D., etc.) [3]	Other [4]			
		4	How many years have you been teaching ESL/EFL at the university level?	(0-5) Years [1]	(6-10) Years [2]	(11-15) Years [3]	(16-20) Years [4]	21 Years or more [5]		
		5	What is the average class size you teach?	10-15 students [1]	16-20 students [2]	21-25 students [3]	26-30 students [4]	26 students or more [5]		
Work-Related Activities		Work-Related Activities During Official Working Hours		1	Assuming a 40-hour work week, please rate the amount of time you spend on [Teaching] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Very little time [1]	Little time [2]	Moderate time [3]	Considerable time [4]	A great deal of time [5]
				2	Assuming a 40-hour work week, please rate the amount of time you spend on [lesson planning] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Very little time [1]	Little time [2]	Moderate time [3]	Considerable time [4]	A great deal of time [5]
				3	Assuming a 40-hour work week, please rate the amount of time you spend on [marking and providing feedback] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Very little time [1]	Little time [2]	Moderate time [3]	Considerable time [4]	A great deal of time [5]
				4	Assuming a 40-hour work week, please rate the amount of time you spend on [administrative tasks] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Very little time [1]	Little time [2]	Moderate time [3]	Considerable time [4]	A great deal of time [5]
				5	Assuming a 40-hour work week, please rate the amount of time you spend on [Other Work-Related Things] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Very little time [1]	Little time [2]	Moderate time [3]	Considerable time [4]	A great deal of time [5]

Teacher Workload and Stress	Work-Related Activities Outside of Official Working Hours	1	On average, how many hours per week do you spend on [Marking student work] OUTSIDE of your regular working hours at college/university?	I do NOT perform any work-related tasks outside of regular working hours [1]	1-3 hours/week [2]	4-7 hours/week [3]	8-12 hours/week [4]	more than 12 hours/week [5]
		2	On average, how many hours per week do you spend on [Writing feedback on assignments] OUTSIDE of your regular working hours at college/university?	I do NOT perform any work-related tasks outside of regular working hours [1]	1-3 hours/week [2]	4-7 hours/week [3]	8-12 hours/week [4]	more than 12 hours/week [5]
		3	On average, how many hours per week do you spend on [Responding to student emails] OUTSIDE of your regular working hours at college/university?	I do NOT perform any work-related tasks outside of regular working hours [1]	1-3 hours/week [2]	4-7 hours/week [3]	8-12 hours/week [4]	more than 12 hours/week [5]
		4	On average, how many hours per week do you spend on [Lesson planning] OUTSIDE of your regular working hours at college/university?	I do NOT perform any work-related tasks outside of regular working hours [1]	1-3 hours/week [2]	4-7 hours/week [3]	8-12 hours/week [4]	more than 12 hours/week [5]
		5	On average, how many hours per week do you spend on [Administrative tasks] OUTSIDE of your regular working hours at college/university?	I do NOT perform any work-related tasks outside of regular working hours [1]	1-3 hours/week [2]	4-7 hours/week [3]	8-12 hours/week [4]	more than 12 hours/week [5]
	Sources of Workload and Stress	3	How often do you feel overwhelmed by the number of students you have to grade?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		2	How much time do you spend on marking assignments and providing feedback per week?	Less than 5 hours [1]	5-10 hours [2]	11-15 hours [3]	16-20 hours [4]	More than 20 hours [5]
		3	To what extent do administrative tasks related to quality assurance standards add to your workload?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		4	How often do you feel that the time you spend on administrative tasks affects your ability to focus on teaching and providing feedback?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
	Impact of Workload on Well-being	1	How often do you feel emotionally drained from your work as an EFL/ESL teacher?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		2	To what extent do you feel that your workload affects your emotional well-being?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		3	How often do you feel that your workload affects your job satisfaction?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		4	How often do you feel overwhelmed by the demands of your teaching responsibilities?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]

Quality Assurance Standards	Impact of Quality Assurance Standards	1	How much additional workload do quality assurance standards add to your teaching responsibilities?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		2	To what extent do quality assurance standards increase your stress levels?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		3	How often do you feel that quality assurance standards limit your ability to provide detailed feedback to your students?	Very ineffectively [1]	Somewhat ineffectively [2]	Neutral [3]	Somewhat effectively [4]	Very effectively [5]
		4	How do you balance maintaining educational quality through standards and managing your workload effectively?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
	Teacher Perspectives on Balancing Standards and Teaching	1	How often do you feel that quality assurance standards are beneficial for improving educational quality?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		2	To what extent do you feel that quality assurance standards are necessary for maintaining educational quality?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		3	How often do you feel that quality assurance standards limit your autonomy and creativity in the classroom?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		4	How challenging do you find it to balance maintaining high educational standards with managing your workload effectively?	Not at all challenging [1]	Slightly challenging [2]	Moderately challenging [3]	Very challenging [4]	Extremely challenging [5]
Written Corrective Feedback	Challenges in Providing Feedback	1	To what extent do large class sizes affect your ability to provide detailed written corrective feedback?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		2	How often do you feel that administrative tasks related to quality assurance standards hinder your ability to provide detailed feedback?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		3	How much time do you typically spend providing detailed written corrective feedback on a single student's writing assignment?	Less than 5 minutes [1]	5-10 minutes [2]	11-15 minutes [3]	16-20 minutes [4]	More than 20 minutes [5]
		4	How often do you feel that the workload related to quality assurance standards and large class sizes affect your emotional well-being?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
	Emotional Impact of Feedback Challenges	1	How often do you feel guilty or stressed about not being able to provide detailed feedback due to workload constraints?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		2	To what extent do you feel that your ability to provide detailed feedback affects your students' learning outcomes?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		3	How often do you feel that your students are negatively impacted by your inability to provide detailed feedback due to workload constraints?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		4	How much do you agree with the statement: "The emotional burden of wanting to provide detailed feedback but being unable to do so affects my overall well-being"?	Strongly disagree [1]	Disagree [2]	Neutral [3]	Agree [4]	Strongly agree [5]

AI-Assisted Tools	Effectiveness of AI Tools	1	To what extent do you believe AI-assisted tools can effectively automate or streamline tasks related to assignment marking and feedback provision?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		2	How often do you think AI-assisted tools could help reduce the time you spend on marking and providing feedback?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		3	To what extent do you trust AI tools to provide fair and unbiased feedback to your students?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		4	How much do you agree with the statement: "AI-assisted tools can help alleviate the workload related to assignment marking and feedback provision"?	Strongly disagree [1]	Disagree [2]	Neutral [3]	Agree [4]	Strongly agree [5]
	Teacher Perceptions of AI Tools	1	How open are you to using AI-assisted tools for marking and providing feedback in your classroom?	Not open at all [1]	Somewhat open [2]	Neutral [3]	Open [4]	Very open [5]
		2	To what extent are you concerned about the potential biases in AI-assisted tools for marking and providing feedback?	Not at all concerned [1]	A little concerned [2]	Somewhat concerned [3]	A lot concerned [4]	Extremely concerned [5]
		3	How much do you think AI tools could enhance the quality of feedback you provide to students?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		4	How do you perceive the potential benefits and challenges of integrating AI-assisted tools into your teaching practices?	Very negatively [1]	Somewhat negatively [2]	Neutral [3]	Somewhat positively [4]	Positively [5]
Innovative Strategies and Technologies	Adopted Strategies for Feedback	1	How often do you use innovative strategies or technologies to facilitate the process of giving written corrective feedback?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		2	How effective do you find these strategies or technologies in managing your workload?	Not effective at all [1]	Somewhat effective [2]	Neutral [3]	Effective [4]	Very effective [5]
		3	How much do these strategies or technologies impact your students' learning outcomes?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		4	How do these strategies or technologies affect your overall well-being as a teacher?	Very negatively [1]	Somewhat negatively [2]	Neutral [3]	Somewhat positively [4]	Positively [5]
	Impact on Teacher Well-being and Student Outcomes	1	How often do these strategies or technologies help you provide more detailed and personalized feedback to your students?	Never [1]	Rarely [2]	Sometimes [3]	Often [4]	Always [5]
		2	To what extent do these strategies or technologies reduce the emotional burden of providing detailed feedback?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		3	How much do these strategies or technologies contribute to your overall job satisfaction?	Not at all [1]	A little [2]	Somewhat [3]	A lot [4]	A great deal [5]
		4	How do your students perceive the effectiveness of the feedback you provide using these strategies or technologies?	Very ineffectively [1]	Somewhat ineffectively [2]	Neutral [3]	Somewhat effectively [4]	Very effectively [5]

Appendix II. The Interview Questions

	Domain		Interview Questions
1	Workload and Burnout	1	Could you describe your typical workweek, including hours spent on teaching, feedback, and administrative tasks?
		2	How has your workload changed over the past few years, particularly regarding documentation requirements?
		3	In what ways does your current workload affect your emotional wellbeing and job satisfaction?"
2	Quality Assurance Impact	1	How do quality assurance requirements affect your daily teaching practices?
		2	What specific QA documentation takes the most time, and how does this impact your ability to focus on teaching?
		3	Do you see a conflict between QA standards and your ability to provide quality feedback? Could you explain?
3	Feedback Challenges	1	With your current class size, how do you approach providing feedback to students?
		2	What challenges do you face in providing detailed feedback to language learners?
		3	Have you developed any strategies to manage the emotional labor of providing critical feedback?
4	AI Perceptions	1	What are your thoughts on using AI tools to assist with administrative tasks or feedback provision?
		2	What concerns do you have about integrating AI into your teaching practice?
		3	How do you think AI tools might affect your professional identity as a language teacher?

Appendix III. Focus Group Discussion Questions

Groups		Focus Group Discussion Questions	
Group 1	Early Career Teachers	1	How do you balance the demands of providing quality feedback with meeting institutional requirements as a new teacher?
		2	What support systems would help you manage workload more effectively?
Group 2	Experienced Teachers	1	How have changes in administrative requirements affected your teaching approach over your career?
		2	What coping strategies have you developed to maintain wellbeing despite increasing demands?
Group 3	Teachers with Large Classes	1	What are the specific challenges of providing feedback in large language classes?
		2	At what class size do you find it becomes significantly more difficult to provide quality feedback?
Group 4	Perspectives on AI	1	What specific tasks would you be most comfortable delegating to AI assistance?
		2	How might AI tools be designed to preserve your professional autonomy while reducing workload?
Group 5	Quality Assurance Standards	1	How could QA frameworks be recalibrated to maintain standards while reducing the administrative burden?
		2	What documentation requirements do you find most valuable versus most burdensome?

Appendix IV. Cronbach Alpha Results for All Questionnaire Questions

Question Number	Question	Construct	Sub-Construct	Reversed ?	Cronbach Alpha
1	Assuming a 40-hour work week, please rate the amount of time you spend on [Teaching] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Work-Related Activities	Work-Related Activities During Official Work Hours	No	0.75
2	Assuming a 40-hour work week, please rate the amount of time you spend on [lesson planning] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Work-Related Activities	Work-Related Activities During Official Work Hours	No	0.45
3	Assuming a 40-hour work week, please rate the amount of time you spend on [marking] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Work-Related Activities	Work-Related Activities During Official Work Hours	No	0.82
4	Assuming a 40-hour work week, please rate the amount of time you spend on [providing feedback] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Work-Related Activities	Work-Related Activities During Official Work Hours	No	0.18
5	Assuming a 40-hour work week, please rate the amount of time you spend on [administrative tasks] in relation to your total working time. Use a scale of 1 to 5, where 1 indicates a very small amount of time and 5 indicates a very large amount of time.	Work-Related Activities	Work-Related Activities During Official Work Hours	No	0.32
6	On average, how many hours per week do you spend on [Marking student work] OUTSIDE of your regular working hours at college/university?	Work-Related Activities	Work-Related Activities Outside Official Work Hours	No	0.8
7	On average, how many hours per week do you spend on [Writing feedback on assignments] OUTSIDE of your regular working hours at college/university?	Work-Related Activities	Work-Related Activities Outside Official Work Hours	No	0.85
8	On average, how many hours per week do you spend on [Responding to student emails] OUTSIDE of your regular working hours at college/university?	Work-Related Activities	Work-Related Activities Outside Official Work Hours	No	0.96
9	On average, how many hours per week do you spend on [Lesson planning] OUTSIDE of your regular working hours at college/university?	Work-Related Activities	Work-Related Activities Outside Official Work Hours	No	0.62
10	On average, how many hours per week do you spend on [Administrative tasks] OUTSIDE of your regular working hours at college/university?	Work-Related Activities	Work-Related Activities Outside Official Work Hours	No	0.9
11	How often do you feel overwhelmed by the number of students you have to grade?	Teacher Workload and Stress	Sources of Workload and Stress	No	0.93
12	How much time do you spend on marking assignments and providing feedback per week?	Teacher Workload and Stress	Sources of Workload and Stress	No	0.78
13	To what extent do administrative tasks related to quality assurance standards add to your workload?	Teacher Workload and Stress	Sources of Workload and Stress	No	0.46
14	How often do you feel that the time you spend on administrative tasks affects your ability to focus on teaching and providing feedback?	Teacher Workload and Stress	Sources of Workload and Stress	No	0.94
15	How often do you feel emotionally drained from your work as an EFL/ESL teacher?	Teacher Workload and Stress	Impact of Workload on Well-being	No	0.99
16	To what extent do you feel that your workload affects your emotional well-being?	Teacher Workload and Stress	Impact of Workload on Well-being	No	0.99
17	How often do you feel that your workload affects your job satisfaction?	Teacher Workload and Stress	Impact of Workload on Well-being	No	0.99
18	How often do you feel overwhelmed by the demands of your teaching responsibilities?	Teacher Workload and Stress	Impact of Workload on Well-being	No	0.99
19	How much additional workload do quality assurance standards add to your teaching responsibilities?	Quality Assurance Standards	Impact of Quality Assurance Standards	No	0.7
20	To what extent do quality assurance standards increase your stress levels?	Quality Assurance Standards	Impact of Quality Assurance Standards	No	0.86
21	How often do you feel that quality assurance standards limit your ability to provide detailed feedback to your students?	Quality Assurance Standards	Impact of Quality Assurance Standards	No	0.87
22	How do you balance maintaining educational quality through standards and managing your workload effectively?	Quality Assurance Standards	Impact of Quality Assurance Standards	No	0.42
23	How often do you feel that quality assurance standards are beneficial for improving educational quality?	Quality Assurance Standards	Teacher Perspectives on Balancing Standards and	No	0.71
24	To what extent do you feel that quality assurance standards are necessary for maintaining educational quality?	Quality Assurance Standards	Teacher Perspectives on Balancing Standards and	No	0.71
25	How often do you feel that quality assurance standards limit your autonomy and creativity in the classroom?	Quality Assurance Standards	Teacher Perspectives on Balancing Standards and	Yes	0.51

26	How challenging do you find it to balance maintaining high educational standards with managing your workload effectively?	Quality Assurance Standards	Teacher Perspectives on Balancing Standards and	No	0.31
27	To what extent do large class sizes affect your ability to provide detailed written corrective feedback?	Written Corrective Feedback	Challenges in Providing Feedback	No	0.6
28	How often do you feel that administrative tasks related to quality assurance standards hinder your ability to provide detailed feedback?	Written Corrective Feedback	Challenges in Providing Feedback	No	0.55
29	How much time do you typically spend providing detailed written corrective feedback on a single student's writing assignment?	Written Corrective Feedback	Challenges in Providing Feedback	Yes	0.14
30	How often do you feel that the workload related to quality assurance standards and large class sizes affect your emotional well-being?	Written Corrective Feedback	Challenges in Providing Feedback	No	0.6
31	How often do you feel guilty or stressed about not being able to provide detailed feedback due to workload constraints?	Written Corrective Feedback	Emotional Impact of Feedback Challenges	No	0.7
32	To what extent do you feel that your ability to provide detailed feedback affects your students' learning outcomes?	Written Corrective Feedback	Emotional Impact of Feedback Challenges	No	0.7
33	How often do you feel that your students are negatively impacted by your inability to provide detailed feedback due to workload constraints?	Written Corrective Feedback	Emotional Impact of Feedback Challenges	No	0.8
34	How much do you agree with the statement: "The emotional burden of wanting to provide detailed feedback but being unable to do so affects my overall well-being"?	Written Corrective Feedback	Emotional Impact of Feedback Challenges	No	0.8
35	To what extent do you believe AI-assisted tools can effectively automate or streamline tasks related to assignment marking and feedback provision?	AI-Assisted Tools	Effectiveness of AI Tools	No	0.93
36	How often do you think AI-assisted tools could help reduce the time you spend on marking and providing feedback?	AI-Assisted Tools	Effectiveness of AI Tools	No	0.99
37	To what extent do you trust AI tools to provide fair and unbiased feedback to your students?	AI-Assisted Tools	Effectiveness of AI Tools	No	0.97
38	How much do you agree with the statement: "AI-assisted tools can help alleviate the workload related to assignment marking and feedback provision"?	AI-Assisted Tools	Effectiveness of AI Tools	No	0.97
39	How open are you to using AI-assisted tools for marking and providing feedback in your classroom?	AI-Assisted Tools	Teacher Perceptions of AI Tools	No	0.85
40	To what extent are you concerned about the potential biases in AI-assisted tools for marking and providing feedback?	AI-Assisted Tools	Teacher Perceptions of AI Tools	Yes	0.33
41	How much do you think AI tools could enhance the quality of feedback you provide to students?	AI-Assisted Tools	Teacher Perceptions of AI Tools	No	0.8
42	How do you perceive the potential benefits and challenges of integrating AI-assisted tools into your teaching practices?	AI-Assisted Tools	Teacher Perceptions of AI Tools	No	0.8
43	How often do you use innovative strategies or technologies to facilitate the process of giving written corrective feedback?	Innovative Strategies and Technologies	Adopted Strategies for Feedback	No	0.99
44	How effective do you find these strategies or technologies in managing your workload?	Innovative Strategies and Technologies	Adopted Strategies for Feedback	No	0.99
45	How much do these strategies or technologies impact your students' learning outcomes?	Innovative Strategies and Technologies	Adopted Strategies for Feedback	No	0.99
46	How do these strategies or technologies affect your overall well-being as a teacher?	Innovative Strategies and Technologies	Adopted Strategies for Feedback	No	0.99
47	How often do these strategies or technologies help you provide more detailed and personalized feedback to your students?	Innovative Strategies and Technologies	Impact on Teacher Well-being and Student Outcomes	No	0.99
48	To what extent do these strategies or technologies reduce the emotional burden of providing detailed feedback?	Innovative Strategies and Technologies	Impact on Teacher Well-being and Student Outcomes	No	0.99
49	How much do these strategies or technologies contribute to your overall job satisfaction?	Innovative Strategies and Technologies	Impact on Teacher Well-being and Student Outcomes	No	0.99
50	How do your students perceive the effectiveness of the feedback you provide using these strategies or technologies?	Innovative Strategies and Technologies	Impact on Teacher Well-being and Student Outcomes	No	0.99