

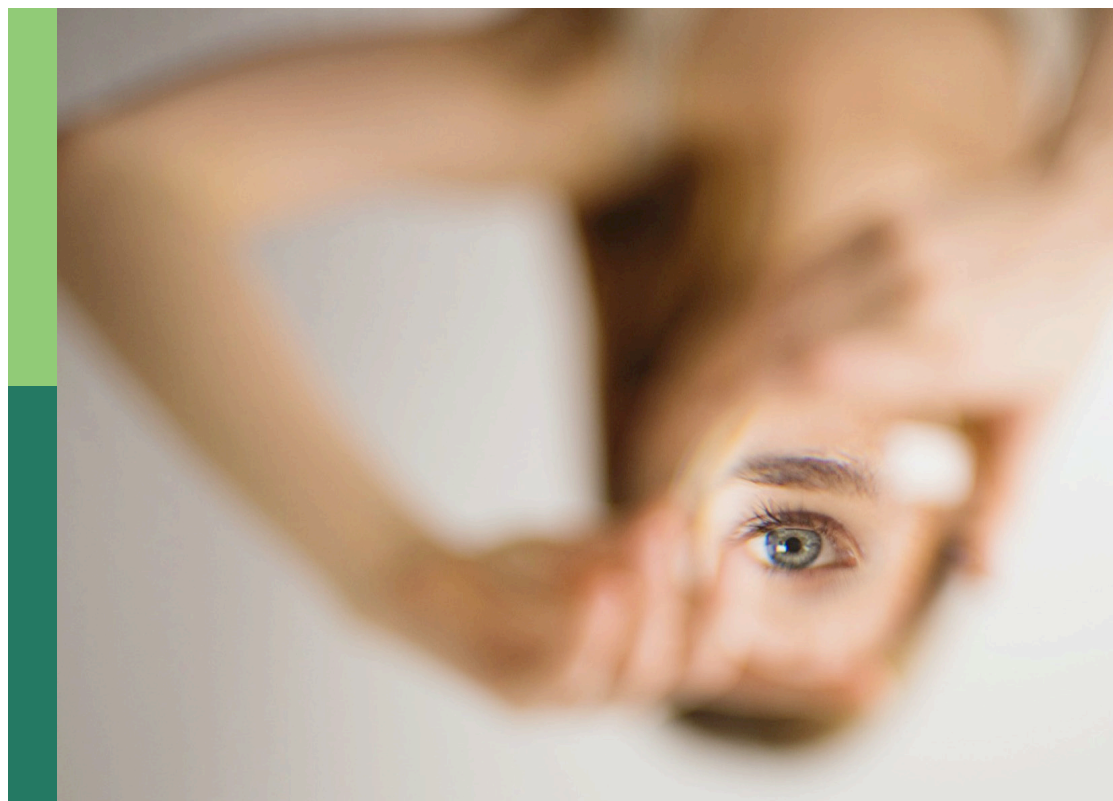
Demystifying academic writing in higher education: a process view on academic textual production

Edited by

Xinghua Liu, Rui Alexandre Alves, Angelique Aitken
and Josef Schmied

Published in

Frontiers in Psychology
Frontiers in Education



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-7000-5
DOI 10.3389/978-2-8325-7000-5

Generative AI statement

Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Demystifying academic writing in higher education: a process view on academic textual production

Topic editors

Xinghua Liu — Shanghai Jiao Tong University, China

Rui Alexandre Alves — University of Porto, Portugal

Angelique Aitken — The Pennsylvania State University (PSU), United States

Josef Schmied — Chemnitz University of Technology, Germany

Citation

Liu, X., Alves, R. A., Aitken, A., Schmied, J., eds. (2025). *Demystifying academic writing in higher education: a process view on academic textual production*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-7000-5

Table of contents

- 04 **Editorial: Demystifying academic writing in higher education: a process view on academic textual production**
Xinghua Liu, Rui A. Alves, A. Angelique Aitken and Josef Schmied
- 06 **Incorporating peer feedback in academic writing: a systematic review of benefits and challenges**
Yuzhu Wei and Donghong Liu
- 20 **A bibliometric study of identity construction in English writing for academic purposes**
Yuhan Tian and Donghong Liu
- 32 **Self-imposed pressure or joyful learning: emotions of Chinese as a foreign language learners in feedback on academic writing**
Ruoxi Liu and Ping Xin
- 46 **Elucidating the emotional persona in the Romanian university students' academic discourse: a corpus-based exploration**
Diana Paula Dudău, Madalina Chitez and Florin Alin Sava
- 67 **EFL learners' motivation and acceptance of using large language models in English academic writing: an extension of the UTAUT model**
Qingran Wang
- 79 **Citation practices in applied linguistics: a comparative study of Chinese expert and novice authors**
Xue Gong, Ruoxi Liu and Chuanbo Ji
- 90 **Four supervisory mentoring practices that support online doctoral students' academic writing**
Sandra Becker, Michele Jacobsen and Sharon Friesen
- 101 **Componential modeling of argumentative essay writing from multiple online sources: a Bayesian network approach**
Anisha Singh, Yuting Sun, Patricia A. Alexander and Hongyang Zhao
- 124 **Teacher motivational strategies in Saudi university EFL writing classes: a qualitative study**
Muhammad M. M. Abdel Latif, Talal Musaed Alghizzi and Tahani Munahi Alshahrani
- 138 **Optimizing academic engagement and mental health through AI: an experimental study on LLM integration in higher education**
Min Zhang



OPEN ACCESS

EDITED AND REVIEWED BY
Daniel H. Robinson,
The University of Texas at Arlington College of
Education, United States

*CORRESPONDENCE
Xinghua Liu
✉ liuxinghua@sjtu.edu.cn

RECEIVED 10 September 2025
ACCEPTED 22 September 2025
PUBLISHED 30 September 2025

CITATION
Liu X, Alves RA, Aitken AA and Schmied J
(2025) Editorial: Demystifying academic
writing in higher education: a process view on
academic textual production.
Front. Psychol. 16:1702487.
doi: 10.3389/fpsyg.2025.1702487

COPYRIGHT
© 2025 Liu, Alves, Aitken and Schmied. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Editorial: Demystifying academic writing in higher education: a process view on academic textual production

Xinghua Liu^{1*}, Rui A. Alves², A. Angelique Aitken³ and
Josef Schmied⁴

¹Shanghai Jiao Tong University, Shanghai, China, ²University of Porto, Porto, Portugal, ³The
Pennsylvania State University (PSU), University Park, PA, United States, ⁴Chemnitz University of
Technology, Chemnitz, Germany

KEYWORDS

academic writing, higher education, process view, teaching and learning of writing, AI

Editorial on the Research Topic

[Demystifying academic writing in higher education: a process view on academic textual production](#)

Writing is not an innate human ability but a skill acquired through training and sustained practice. Nonetheless, it plays a central role in how students learn and how their learning is assessed. Academic writing, where students communicate scholarly ideas, presents unique challenges. At the higher education level, strong academic writing requires clarity of reasoning, mastery of disciplinary knowledge, and linguistic proficiency.

Decades of research have deepened our understanding of the textual features of students' academic writing and the cognitive processes involved. However, a comprehensive process-oriented perspective on academic writing remains underdeveloped. Therefore, the purpose of this Topical Research is to explore how higher education students generate ideas, draft their texts, utilize technology, sustain academic integrity, and finalize their written work.

This Research Topic features four contributions examining the linguistic features of academic writing. A key characteristic of successful academic writing is the construction of authorial identity. [Tian and Liu's](#) systematic review reveals that over the past three decades, the most prominent Research Topics have centered on plagiarism/academic integrity and sociocultural perspectives on identity construction. Their findings hold particular relevance in the current AI-driven era, where the use of AI writing tools has become ubiquitous.

Through corpus-based linguistic analysis, [Dudău et al.](#) identified distinct emotional patterns in Romanian vs. English academic writing. They found that Romanian texts consistently exhibited greater formality and indirectness, which they believe have been shaped by language, cultural norms, and academic conventions.

Gong et al. investigated citation practices among novice and expert authors in the field of Chinese Applied Linguistics. Their study found striking similarities between the two groups, with minimal cross-linguistic differences in citation practices between English and Chinese academic writing.

By using a Bayesian network approach, Singh et al. modeled the cognitive processes of argumentation. Their research highlights students' primary challenges during argumentative writing, namely the framing of counterarguments and the development of in-depth and critical analyses of problems.

Feedback is an integral component of teaching and learning writing, and this Research Topic includes two papers addressing this topic. Liu and Xin explored the emotional responses of Chinese as a Foreign Language (CFL) learners when receiving oral and written feedback from their teachers, and examined their emotion regulation strategies as well. Their findings reveal that teacher feedback elicits three types of emotions: academic achievement emotions, cognitive emotions, and social emotions. The study also highlights that students employ three primary strategies to manage negative emotions: emotion-oriented, appraisal-oriented, and situation-oriented approaches. Wei and Liu conducted a systematic review of peer feedback research in academic writing from 2014 to 2024. They identified five key benefits of peer feedback activities: affective, cognitive, behavioral, social, and meta-cognitive benefits. Additionally, they pinpointed three major challenges associated with peer feedback: difficulties arising from feedback providers, receivers, and contextual factors.

Writing is a cognitively demanding task and presents additional challenges for students learning to write in a foreign language. Therefore, cultivating and sustaining students' motivation is crucial for the success of academic writing instruction. This Research Topic includes two studies on writing motivation. Abdel Latif et al. surveyed experienced English writing teachers from five Saudi universities, identifying eight effective motivational strategies, such as optimizing teacher feedback, negotiating writing topic choices. Their findings also suggest that smaller class sizes facilitate the implementation of these strategies. For doctoral students, mastering academic writing is particularly critical, as it serves as the primary gateway to the academic community. Becker et al. employed a comparative case study approach to examine online mentoring dynamics. They identified five key factors that can help build trust and collaboration between supervisors and research students.

The pervasive influence of AI has made it imperative to integrate AI technology into academic writing instruction. This Research Topic includes two relevant contributions on this topic. Zhang's study contributes to the growing evidence supporting AI-assisted writing instruction. The research demonstrates that when AI tools are used in a guided, structured manner, university students report improved writing quality, enhanced perceived mental wellbeing, and greater academic engagement. Wang's questionnaire survey of Chinese EFL students reveals key insights into their use of large language models (LLMs) for business English writing. The findings indicate that performance expectancy and social influence strongly predict students' intention to use

LLMs. What is particularly interesting is that motivation not only influences students' perception of the usefulness of LLMs, but also determines students' actual use of them in their writing processes.

The papers in this Research Topic explore diverse aspects of academic writing in higher education, demonstrating both the richness and complexity of this field. While significant progress has been made, we identify three critical areas requiring further investigation. First, the planning, composing, and revision stages of academic writing remain largely unexplored. A deeper understanding of students' challenges and effective instructional strategies is urgently needed. Second, as AI is transforming education at all levels, it is imperative to study how we can maximize AI's benefits for teaching and learning writing while mitigating its potential risks and ethical concerns. Third, we believe good academic writing instruction needs to help develop autonomous writers. While more research is warranted for developing students' self-regulation skills and equipping them with the independence needed for lifelong academic success, the articles collected here are already insightful pointers in that direction.

Author contributions

XL: Writing – original draft, Writing – review & editing. RA: Writing – original draft, Writing – review & editing. AA: Writing – original draft, Writing – review & editing. JS: Writing – original draft, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



OPEN ACCESS

EDITED BY

Xinghua Liu,
Shanghai Jiao Tong University, China

REVIEWED BY

Ying Gao,
Northeast Normal University, China
Marianne Ellegaard,
University College Copenhagen, Denmark
Shengkai Yin,
Shanghai Jiao Tong University, China
Xiaodong Zhang,
Beijing Foreign Studies University, China
Baohua Dong,
Chongqing University of Science and
Technology, China
Jinlong Yang,
Nanjing University, China
Doaa Hamam,
Higher Colleges of Technology,
United Arab Emirates

*CORRESPONDENCE

Donghong Liu
✉ liudonghong19@hotmail.com

RECEIVED 06 October 2024

ACCEPTED 04 November 2024

PUBLISHED 26 November 2024

CITATION

Wei Y and Liu D (2024) Incorporating peer
feedback in academic writing: a systematic
review of benefits and challenges.
Front. Psychol. 15:1506725.
doi: 10.3389/fpsyg.2024.1506725

COPYRIGHT

© 2024 Wei and Liu. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Incorporating peer feedback in academic writing: a systematic review of benefits and challenges

Yuzhu Wei and Donghong Liu*

School of Foreign Languages, Southeast University, Nanjing, China

Academic writing is paramount to students' academic success in higher education. Given the widely acknowledged benefits of peer feedback in diverse learning contexts, such as fostering a positive psychological mindset, there has been a growing interest in applying this approach to facilitate the development of academic writing. This study is launched to examine the primary features and findings of the studies that have investigated the benefits and challenges of the utilization of peer feedback in academic writing development. The methodology of this study incorporates a rigorous literature search methodology, encompassing database search, reference search, and manual search, which is subsequently followed by a content analysis of the selected studies. With the guidance of PRISMA 2020, a total of 60 related articles, spanning the period from 2014 to 2024, are selected through title screening, abstract screening and content screening, adhering to strict inclusion and exclusion criteria. The findings of this study reveal a growing global interest in peer feedback in academic writing, and highlight the need for future research on masters'/doctoral students and quantitative approaches to deepen understanding of its effects. Moreover, 16 distinct benefits of peer feedback in the academic writing context were delineated and subsequently categorized into five categories: affective benefits, cognitive benefits, behavioral benefits, social benefits, and meta-cognitive benefits. Furthermore, an analysis of the implementation challenges revealed 13 types of obstacles, which were traced to three primary sources: challenges originating from feedback receivers, those posed by feedback providers, and those stemming from the peer feedback settings. Based on these findings, several pedagogical and future research suggestions are proffered to guide both the practitioners and researchers.

KEYWORDS

academic writing, peer feedback, affective benefits, benefits and challenges, psychological mindset

1 Introduction

In higher education, academic writing is considered a core competency for students (Chakraborty et al., 2021). Effective academic writing as the currency of intellectual exchange, which facilitates the sharing of novel insights and contributes to the advancement of knowledge, is crucial for the students' academic success and career development (Aitchison and Lee, 2006; Swales and Peak, 1994). To date, significant emphasis has been placed on the academic writing instruction (Schillings et al., 2023). Defined as a process whereby students critically assess the level, merit, or quality of their peers' work (Topping, 2009), peer feedback has garnered significant attention in recent years as an active learning strategy that fosters interaction, collaboration, and reciprocal learning (Liu and Carless, 2006). The integration of peer feedback into academic development is underpinned by theoretical frameworks that emphasize the social nature of learning and the role of collaborative interactions in the

development of cognitive and metacognitive skills, such as the Collaborative Learning Theory in social psychology Bruffee (1984) and Vygotsky's (1978) sociocultural theory.

In recent years, there has been a surge of interest in the role of peer feedback in academic writing education. Numerous studies have validated the effectiveness of these diverse peer feedback practices in advancing academic writing development. Prominently, peer feedback serves as a catalyst for elevating students' academic writing quality and refining their academic writing skills (Kostopoulou and O'Dwyer, 2021; Rodas and Colombo, 2021). Beyond this, by engaging students in the evaluation process, peer feedback fosters a deeper understanding of academic writing criteria, promotes self-reflection, and enhances critical and analytical skills (Boillos, 2024; Davis, 2014; Kostopoulou and O'Dwyer, 2021; Osman et al., 2022), empowering students to become more discerning consumers and producers of academic texts (Ciampa and Wolfe, 2023; Pugh and Veitch, 2019; Yu, 2019). Furthermore, the collaborative nature of peer feedback encourages a sense of academic community and belonging within the learning environment, which can positively impact students' motivation and engagement in the writing process (Geithner and Pollastro, 2016; Yallop et al., 2021).

However, the implementation of peer feedback in the academic writing context is not without its challenges. Insufficient feedback proficiency and domain-specific knowledge often translate into unproductive and unreliable feedback (Ciampa and Wolfe, 2023; Colombo and Rodas, 2021; Kostopoulou and O'Dwyer, 2021; López-Pellisa et al., 2021; Xu and Zhang, 2023). Moreover, the potential for interpersonal friction arising from the delivery of critical feedback and the risk of providing inadequate feedback pose further obstacles (Cheong et al., 2023; Ciampa and Wolfe, 2023; Rodas and Colombo, 2021; Yu, 2021). Some students may experience anxiety and insecurity when engaged in peer feedback activities, as they highly value camaraderie and harmony within their reviewing group (Xu and Li, 2018; Xue et al., 2023). Additionally, given the complexity of academic writing, learners tend to harbor a lower level of trust in peer feedback, particularly when juxtaposed against instructor-led feedback, underscoring the need for strategic interventions to address these concerns (Eppler et al., 2021; Pugh and Veitch, 2019; Xu and Zhang, 2023).

Despite the significant contributions of prior research in elucidating the merits and obstacles associated with integrating peer feedback into academic writing, a notable limitation persists in that these studies have focused narrowly on isolated facets of these benefits and challenges. A comprehensive synthesis of the broader spectrum of benefits and challenges has not been realized. Given the complexities of identified challenges, a comprehensive understanding of the potential challenges associated with the implementation of peer feedback in academic writing is conducive to effectively leveraging its advantages in practical applications. Furthermore, acknowledging the heterogeneous nature of benefits and challenges as identified in prior research, there is a compelling need for a systematic synthesis and taxonomy. Such an endeavor would significantly enrich our understanding and inform both instructional strategies and future research endeavors in this domain.

In response to this research gap, the present systematic literature review aims to provide a systematic synthesis of the empirical evidence on the benefits and challenges of incorporating peer feedback into academic writing instruction. Additionally, it aspires to discern trends

in this realm, thereby offering guidance to both practitioners and researchers alike. To achieve this, this review will be guided by the following research questions:

- 1 What are the primary features of contemporary research investigating the effects of peer feedback in academic writing?
- 2 What are the multifaceted benefits of incorporating peer feedback into academic writing education, and how do they contribute to student learning and development?
- 3 What are the primary challenges encountered in implementing peer feedback in academic writing, and how do they affect the feedback process and its outcomes?

2 Methods

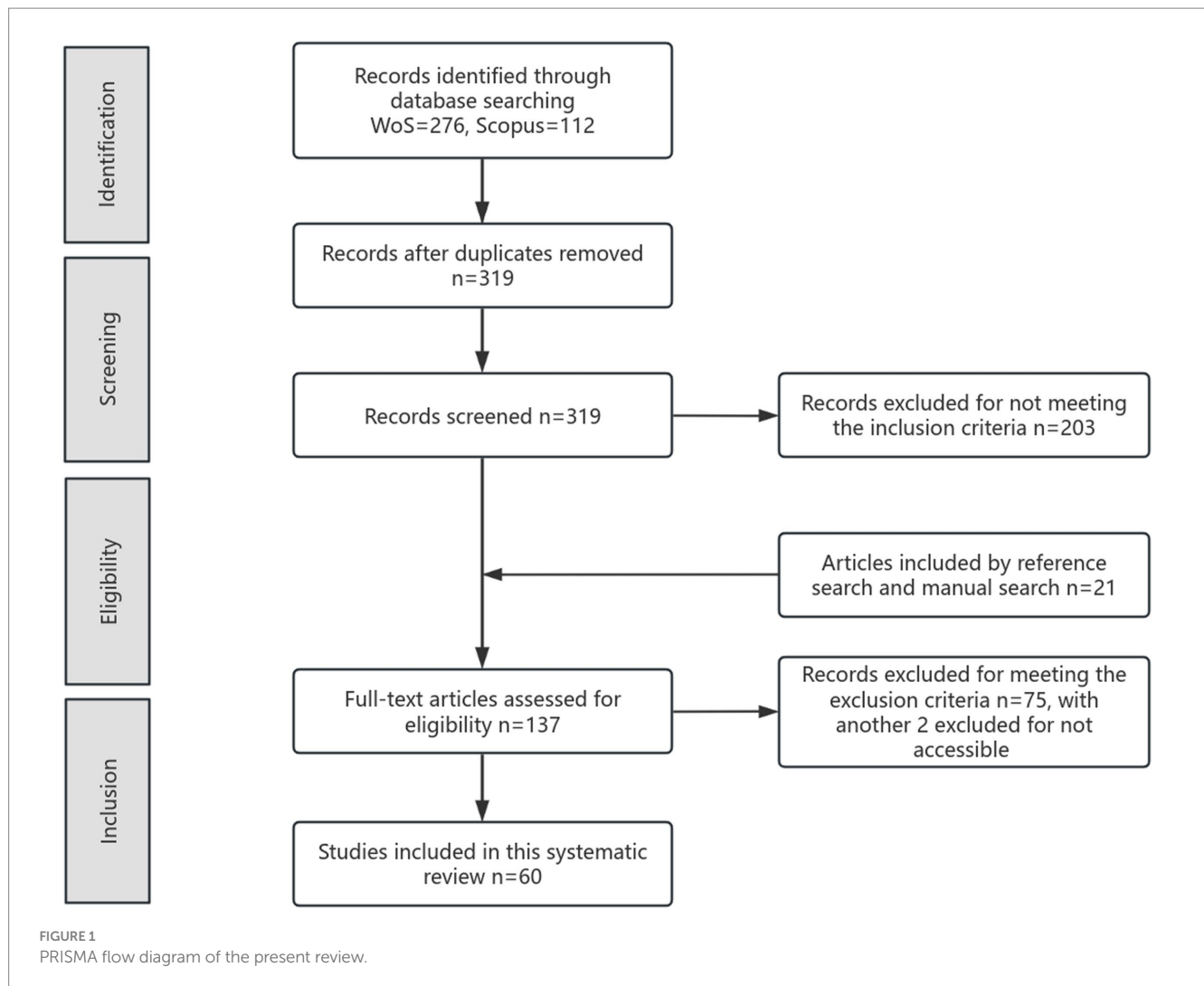
This study employed a systematic review methodology which entails a systematic collection and synthesis of pertinent articles guided by specific research questions (Aromataris and Pearson, 2014; Pearson, 2004; Siddaway et al., 2019). This approach allows researchers to produce more comprehensive and reliable conclusions by integrating diverse findings from previous studies, thereby providing insights for further research and practical applications. Though it was originally developed in medical sciences (Chalmers et al., 2002), numerous studies in the field of education have also attested to its effectiveness and utility (Andrews and Harlen, 2006; Bearman et al., 2012; Davies, 2000; Martin et al., 2020).

2.1 Data collection

To guarantee the credibility of findings, this systematic review followed the guidance of PRISMA 2020 statement (Page et al., 2021), incorporating four stages of data collection: identification, screening, eligibility and inclusion. Details of this procedure are displayed in Figure 1.

In the identification of pertinent studies, three search methods were implemented: a database search, a reference search and a manual search. For the database search, Web of Science Core Collection and Scopus selected as sources owing to their esteemed reputation for encompassing extensive and high-caliber educational research. As a supplementary approach, the reference search was conducted to augment the search process by examining the cited references within the selected studies, thereby mitigating the risk of overlooking significant research contributions. Furthermore, a manual search was conducted utilizing Google Scholar as a platform to identify additional scholarly articles pertaining to the same subject matter.

Prior to embarking on a search for the relevant papers, index terms for the two main concepts, "academic writing" and "peer feedback," were determined by inspecting search terms in previous review studies (Huisman et al., 2019; Yu and Lee, 2016; Zheng et al., 2019), and terminologies used in seminal and recent academic literature. This process resulted in 11 terms for "peer feedback": peer assessment, peer feedback, peer review, peer evaluation, peer rating, peer scoring, peer grading, peer editing, peer response, peer interaction and student feedback, and three terms for "academic writing": academic writing, research writing and scientific writing.



These English terms were used in the search of relevant studies in Web of Science Core Collection, Scopus, and Google Scholar.

Regarding the concept of academic writing, the current study adopts the definition provided by Hyland (2004) and Swales (1990), which posits that academic writing constitutes the formal communication of research and ideas within a specific discipline, adhering to established conventions to contribute to and engage with the field's knowledge. Therefore, this study focuses on various writing genres that differ from school writing, such as course essays, project reports, research proposals, lab notes, journal articles, conference papers, theses, and dissertations, as part of academic writing, regardless of whether they are written in a first or second language. The first author conducted the database search in July 2024, during which only peer-reviewed empirical studies published after 2013 were included for further examination. The initial literature search identified 276 articles from Web of Science Core Collection and 112 from Scopus. After removing 69 duplicates, 319 articles were selected for title and abstract screening to examine whether they meet the inclusion criteria. The following inclusion criteria were used to ensure the relevance and quality of selected articles: (1) published between 2014 and 2024; (2) empirical research; (3) articles concerning peer feedback to academic writing in higher education; (4) articles written in English. After that,

21 relevant papers identified by reference search and manual search were added to the results, which formed a refined pool of 340 articles for eligibility test through full text analysis. It was conducted under the guidance of following exclusion criteria: (1) articles not concerning peer feedback in higher education; (2) articles not revealing the benefits or the problems of peer feedback; (3) articles not clearly demonstrating the context of academic writing; (4) articles not involving peer feedback on their peer's academic writing. Ultimately, this rigorous selection process yielded a total of 60 peer-reviewed empirical studies which were deemed most pertinent for investigating the multifaceted benefits and problems of peer feedback within the academic writing context. This process was visualized in Figure 1.

2.2 Data analysis

The present study employed a conventional content analysis (Hsieh and Shannon, 2005) to delve into the primary features of studies examining the effects of peer feedback in an academic writing context, as well as to identify and analyze the specific benefits and challenges that have been discerned. Conventional content analysis is an inductive category development method during which researchers

TABLE 1 Charting categories, subcategories, and description.

Categories	Subcategories	Description
Descriptive information	Year of publication	It refers to the year when the article was officially published in print, except in cases where it is exclusively published digitally.
	Country of article	It refers to country where the study was conducted.
	Research methodology	It refers to the methodology used in the article, including qualitative, quantitative, and mixed methodology.
	Sample group	It refers to the educational level of the participants, such as undergraduate students, master's students, and doctoral students.
	Subject domain	It refers to the subject domain of the academic writing, such as natural science, social science, and engineering and technological science.
	Task types	It refers to the specific genre of academic writing, such as the scientific paper, scientific report, and research proposal.
Findings	Benefits	It refers to the benefits of incorporating peer feedback in academic writing.
	Challenges	It refers to the challenges encountered in incorporating peer feedback in academic writing.

refrain from relying on pre-determined categories and instead engage in an immersive process of data exploration, from which categories are derived (Kondracki and Wellman, 2002). It is typically deemed suitable in scenarios where the existing theoretical framework or research literature pertaining to a particular phenomenon is limited (Hsieh and Shannon, 2005), aligning well with the circumstances of the current study.

The collected data were scrutinized with the aim of analyzing the overarching themes and discernible patterns of the findings in the selected literature (Opfer and Pedder, 2011). Specifically, Microsoft Excel was employed to record and analyze the codes. To identify the primary features of the studies into the effects of peer feedback in academic writing, subcategories of descriptive information of the reviewed articles were analyzed. This encompassed an examination of the temporal distribution of reviewed studies by year, the research methodologies employed, the educational level of the participants, the subject domain of the academic writing, the task types involved, as well as the geographical locations and educational contexts within which these studies were undertaken (see Table 1). Moreover, contents regarding benefits and challenges of peer feedback underwent a three-stage analytical process. In the first stage, articles were coded using the words in the original text. As the author progressed through the data analysis, efforts were made to minimize the introduction of new codes, giving precedence to existing codes unless novel data emerged that could not be accommodated by them. Following the completion of coding all articles, a meticulous review of the data within each specific code was undertaken to explore potential combinations and segregations, which leads to the formation of distinct subcategories pertaining to benefits and challenges. To ensure reliability, the first author conducted two rounds of coding on all articles, with a two-month interval between the two coding sessions. The consistency rate of coding across all subcategories was not less than 93.3%, indicating a high level of reliability.

3 Findings

3.1 Primary features of the reviewed articles

Figure 2 depicts the annual temporal distribution of the reviewed studies, illustrating the evolution of research endeavors over time.

Prior to 2014, scholarly investigation into the effects of peer feedback within the context of academic writing had already emerged. Despite fluctuations observed in the escalating engagement within this domain, the past 3 years have exhibited a heightened level of interest compared to previous years. Given that the data was collected in July 2024, it is anticipated that the count of related articles for the year 2024 will surpass nine, signifying a substantial growth trend.

The methodologies employed across the reviewed literature were scrutinized. It was found that the mixed methodology emerges as the most prevalent approach ($n = 24$), closely followed by the qualitative methodology ($n = 22$). Conversely, the quantitative methodology is the least utilized ($n = 14$).

Moreover, the reviewed articles have investigated the effects of peer feedback in academic writing context utilizing data sourced from participants with various educational levels, such as undergraduates, master's students and doctoral students. Predominantly, these studies have focused on examining the effects of peer feedback on undergraduates' academic writing ($n = 30$), comprising approximately 50% of the reviewed corpus. Comparable emphasis has been placed on master's students ($n = 18$) and doctoral students ($n = 19$), with a marginal increase in attention toward the latter. Moreover, a subset of studies ($n = 6$) has extended its scope to include participants from alternative educational levels, for instance, pre-master and pre-bachelor programs. It is pertinent to clarify that when studies encompass participants spanning multiple educational levels, they are accounted for within each respective subgroup, thereby leading to a cumulative total of subgroups exceeding the overall count of reviewed articles. The same calculating method is employed in the examination of subject domain, task type and country.

The subject domains and task types of the academic writing in the reviewed articles are visually depicted in Figures 3, 4 respectively. The utilization of peer feedback as a strategy in the development of academic writing has been observed across a diverse spectrum of subject domains. Notably, this approach was the most prevalent in the humanities and social sciences ($n = 31$), significantly outnumbering its application in natural sciences ($n = 10$) and engineering and technological sciences ($n = 7$), which occupy the second and third positions, respectively. Marginal but noteworthy attention was also accorded to the academic writing context within mathematics ($n = 2$), health sciences ($n = 2$), and art and design sciences ($n = 1$). Furthermore, an additional 13 articles existed that did not explicitly delineate the subject domain within

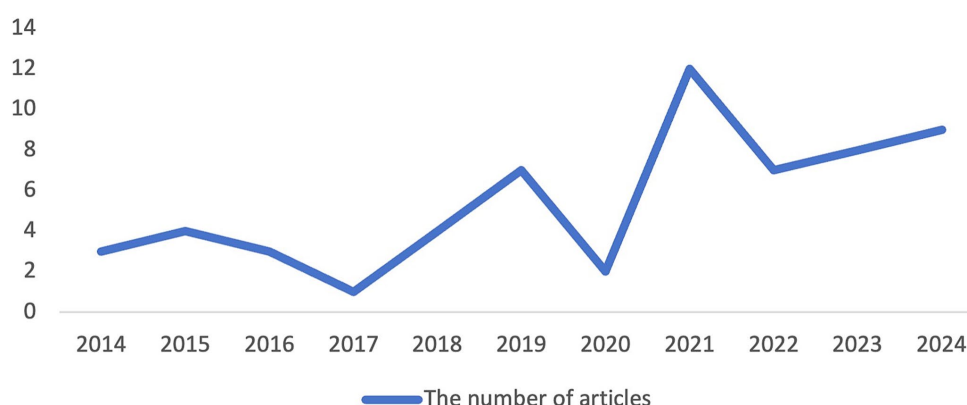


FIGURE 2
Temporal distribution of reviewed articles by year.

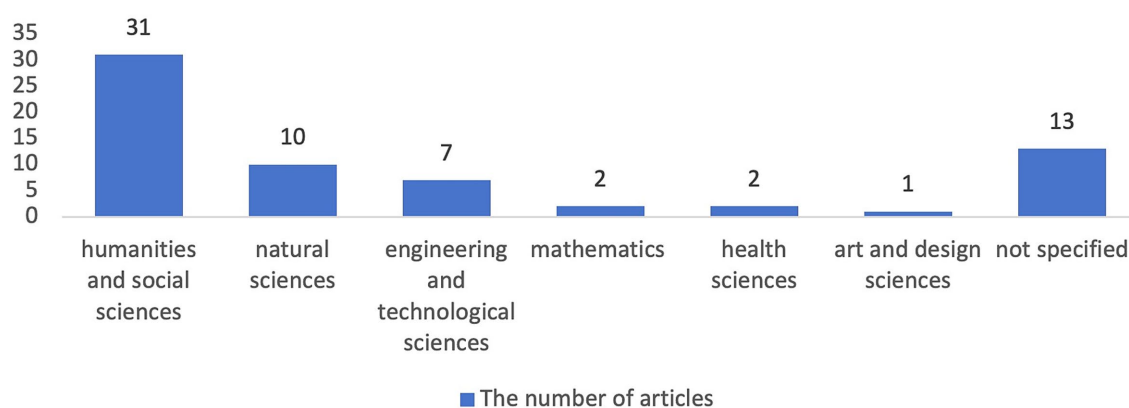


FIGURE 3
Subject domain of the academic writing.

which academic writing was being addressed. Regarding the distribution of task types, scientific papers constitute the most frequently encountered academic writing assignment within the reviewed articles ($n = 16$), followed by scientific reports, which represent a much smaller proportion ($n = 7$). Beyond these general categories, a notable number of studies delved into the writing of specific components within a scientific paper, specifically abstracts ($n = 6$), introductions ($n = 4$), methodologies ($n = 2$), and literature reviews ($n = 6$). Furthermore, the scope of academic writing examined also encompassed thesis/dissertation-related works, more specifically thesis/dissertation proposals ($n = 2$) and thesis drafts ($n = 6$). Additionally, research proposals ($n = 5$) and course essays ($n = 6$) also received similar attention.

Figure 5 delineates the countries (regions) and educational contexts within which these investigations into the integration of peer feedback in academic writing development were conducted. This strategy was observed to be embraced across a diverse range of countries and regions ($n = 20$), underscoring its widespread popularity in the realm of academic writing instruction. Notably, the preponderance of related research was situated in the United States ($n = 12$) and within China, encompassing mainland China ($n = 10$), Macau ($n = 6$), and Hong Kong ($n = 5$), collectively

accounting for 55% of the total reviewed articles. In terms of educational contexts, three distinct modalities were identified for the implementation of peer feedback: courses, workshops, and other informal settings, such as self-organized writing groups. A dominant proportion of the studies were carried out within the structured environment of formal courses ($n = 48$, 81.7% of the total), with a minority being conducted in workshop settings ($n = 6$) and within informal contexts ($n = 5$). Notably, the examination of peer feedback's effects in workshop settings was confined to a limited number of countries, including the United States ($n = 3$), Spain ($n = 1$), Argentina ($n = 1$), and Syria ($n = 1$). Conversely, no studies examining peer feedback in workshop contexts were found to have been conducted in China.

3.2 Specific benefits and challenges identified in previous studies

After conducting a content analysis of the reviewed literature, this study uncovered both the benefits and challenges associated with integrating peer feedback into academic writing. These findings are systematically organized and presented in Tables 2, 3.

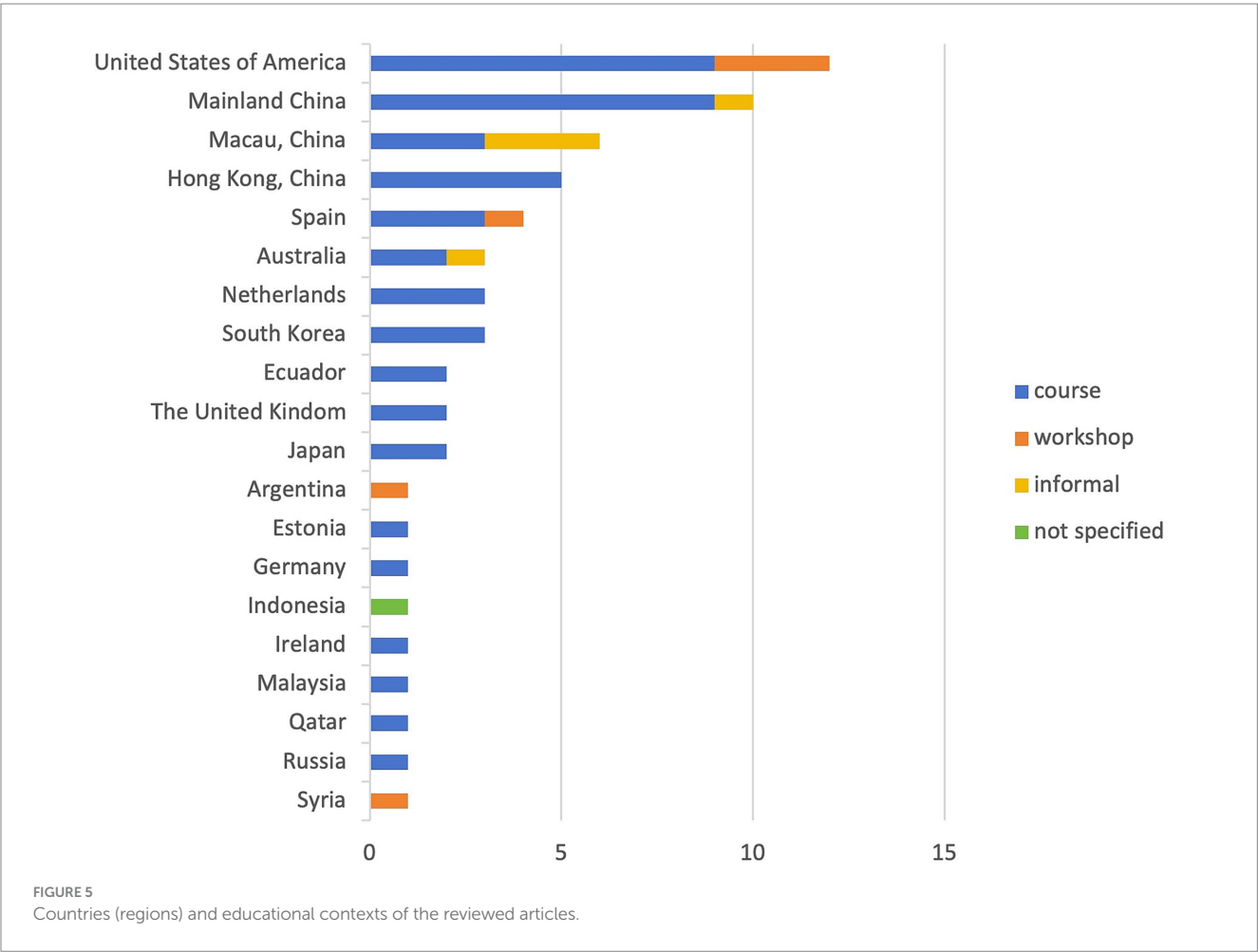
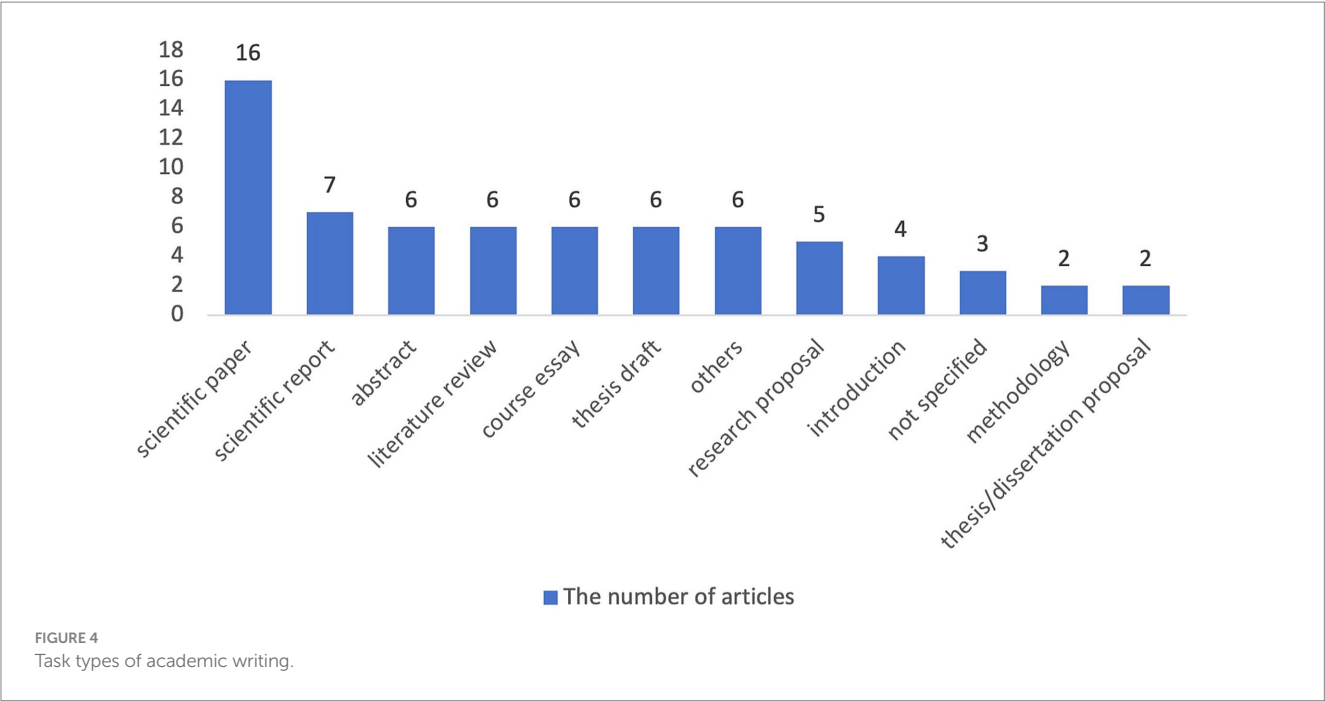


TABLE 2 Benefits of incorporating peer feedback in academic writing.

Categories of benefits	Subcategories of benefits	Number of articles	Sample article
Cognitive benefits	Improving critical and analytical skills	15	Osman et al. (2022)
	Improving academic writing skills	11	Ramon-Casas et al. (2019)
	Knowing more about peer review process	6	Eppler et al. (2021)
	Developing communication skills	5	Gumusoglu et al. (2022)
	Developing feedback literacy	3	Wu and Lei (2023)
	Strengthening subject knowledge	1	Goh et al. (2019)
Behavioral benefits	Improving writing quality	33	Shulgina et al. (2024a)
Affective benefits	Strengthening confidence in academic writing	9	Xu and Zhang (2023)
	Strengthening confidence in critically analyzing academic work	5	Davis (2014)
	Improving willingness to ask for help in the future	1	Liu et al. (2021)
	Increasing motivation toward academic writing	1	Yallop et al. (2021)
Social benefits	Constructing academic community	7	Man et al. (2018)
	Gaining social support	4	Santelmann et al. (2018)
	Strengthening interpersonal relationship	2	Liu et al. (2021)
Meta-cognitive benefits	Promoting self-reflection	16	Yu (2019)
	Increasing metacognitive awareness of the writing process	9	Santelmann et al. (2018)

TABLE 3 Challenges in incorporating peer feedback in academic writing.

Categories of challenges	Subcategories of challenges	Number of articles	Sample article
Challenges from peer feedback receivers	Inadequate feedback literacy	7	Wu and Lei (2023)
	Negative attitude to peer feedback	5	Álvarez et al. (2015)
	Individual difference in gaining benefits	4	Ramon-Casas et al. (2019)
	Heavy cognitive load	2	Shulgina et al. (2024b)
	Low text quality	1	Pugh and Veitch (2019)
	Dependence on peer feedback	1	Lu et al. (2023)
Challenges from peer feedback providers	Students' deficiency in providing constructive feedback	21	Cheong et al. (2023)
	Lack of confidence in providing constructive feedback	6	Ciampa and Wolfe (2023)
	Disregarding providing feedback	5	Xu and Li (2018)
Challenges from peer feedback settings	Interpersonal concerns	8	Zhang et al. (2022)
	Problems of distracting factors	2	Ahmed and Al-Kadi (2021)
	Ineffective grouping of peers	2	Ahmed (2021)
	High time demand	1	Ahmed (2021)

The benefits, in particular, were classified into five distinct categories, each corresponding to a specific facet of enhancement: cognitive, behavioral, affective, social, and meta-cognitive benefits. Cognitive benefits pertain to the development of intellectual abilities such as thinking, knowledge representation, information processing, and decision-making, which are essential for the construction of knowledge during the learning process (Liu et al., 2022; Potvin et al., 2018; Svalberg, 2009; Swain, 2013). Behavioral benefits are associated with the positive changes in students' external actions and academic activities (Nazamud-din et al., 2020; Uher, 2016). Affective benefits relate to the positive influence of peer feedback on students' emotional experiences, including their confidence, willingness, and motivation (Gondim and Mutti, 2011;

Nazamud-din et al., 2020; Piaget, 1962). Social benefits are linked to the positive effects of peer feedback on student interactions within the context of language learning (Svalberg, 2009). Lastly, metacognitive benefits involve the enhancement of self-reflection and the ability to regulate cognition, which are critical for optimizing learning (Goupil and Kouider, 2019; Moses and Baird, 1999). In the reviewed articles, self-reflection and metacognitive awareness of the writing process, such as audience awareness and writer awareness, was found to be improved by the peer feedback practice.

The challenges of incorporating peer feedback in academic writing are found to have three sources: challenges from peer feedback receivers, challenges from peer feedback providers and challenges

from peer feedback settings. Subcategories of them are presented in Table 3.

4 Discussion

4.1 Primary features of the research examining the effects of peer feedback in academic writing

The visualization of the temporal distribution of the reviewed articles exhibits a nuanced dynamic, wherein despite fluctuations in the annual count of related studies, a general trend of escalating interest in the subject matter is discernible over the years. Notably, the past 3 years have witnessed a sustained increase in the level of engagement with this topic. This can be attributed to the heightened emphasis accorded to academic writing within higher education (Chakraborty et al., 2021), and the burgeoning popularity of peer feedback mechanisms in academic writing development, along with their acknowledged merits (Boillos, 2024; Kostopoulou and O'Dwyer, 2021; Osman et al., 2022).

In terms of the research methodologies adopted within the corpus of reviewed articles, a comparative analysis reveals that mixed-methods and qualitative approaches occupy comparable and substantial proportions (40 and 37%, respectively), whereas quantitative methods are less prevalent, accounting for merely 23% (14 studies). Furthermore, despite over half of the reviewed studies incorporating quantitative data analysis, a closer inspection reveals that the majority of these studies focused exclusively on assessing the influence on enhancing writing quality, leading to a notable absence of quantitative data pertaining to other facets. This disparity underscores the need for a more robust quantitative interrogation to validate the discerned benefits and obstacles, thereby fostering a deeper understanding of the topic.

Regarding the educational levels of participants, half of the studies examined undergraduates, whereas doctoral and master's students were involved in 32 and 30% of the research samples, respectively. This preponderance of undergraduate focus likely stems from the heightened importance attributed to academic writing instruction at the undergraduate level within higher education systems. Conventionally, it is assumed that postgraduates, having completed their undergraduate studies, possess a foundational proficiency in academic writing (Sallee et al., 2011; Singleton-Jackson and Lumsden, 2009). However, research findings challenge this notion, revealing that post-graduates often encounter challenges in academic writing and continue to require instructional support (Santelmann et al., 2018; Kabaran, 2022). In light of this revelation, further investigations are imperative to delve into the effectiveness of peer feedback mechanisms for postgraduate students. Such studies would not only elucidate the specific impact of peer feedback on enhancing postgraduate academic writing but also facilitate the strategic integration of this method into the development of postgraduate writing competencies, ultimately contributing to the holistic advancement of academic writing skills across all levels of higher education.

In terms of the subject domains of academic writing, a discernible hierarchy emerges, with humanities and social sciences ($n=31$) occupying the foremost position, followed by natural sciences ($n=10$), and engineering and technological sciences ($n=7$). Conversely,

mathematics ($n=2$), health sciences ($n=2$), and art and design sciences ($n=1$) received comparatively limited attention. This distribution may be attributed to the substantial student enrollment in the aforementioned major disciplines, along with their relatively greater accessibility. Furthermore, an analysis of the task types in the reviewed articles reveals a predilection toward scientific papers and their constituent elements. Specifically, the abstract, introduction, literature review, and methodology garnered exceptional emphasis, likely stemming from their pivotal role in shaping the integrity and rigor of a scientific paper. This underscores the criticality of these components in contributing to the overall quality and comprehension of scientific research.

The analysis of national landscapes and educational contexts within the reviewed articles underscores the widespread adoption of peer feedback as a strategy for enhancing academic writing capabilities across diverse countries, spanning from the United States of America to Syria. This trend underscores the popularity and efficacy of peer feedback in fostering academic writing development (Kostopoulou and O'Dwyer, 2021; Rodas and Colombo, 2021). The United States leads the way in research endeavors, with the highest number of studies conducted ($n=12$), closely followed by mainland China ($n=10$), Macau China ($n=6$), and Hong Kong China ($n=5$). This distribution indicates that both the United States and China prioritize peer feedback as a vital tool in nurturing academic writing skills. However, it is noteworthy that the majority of the reviewed studies (82%) implemented peer feedback within the confines of formal coursework. Despite China's significant contribution to the field, notably absent are studies examining peer feedback in a workshop setting. This observation may be attributed to the fact that academic writing instruction is predominantly conducted within classroom environments, whereas workshop organizers may not fully recognize the inherent value of peer feedback as an instructional method.

4.2 Benefits and challenges identified in incorporation peer feedback in academic writing

In the examination of the reported benefits of incorporating peer feedback in academic writing, various benefits have been identified which can be divided into five categories, namely, cognitive benefits, behavioral benefits, affective benefits, social benefits, and meta-cognitive benefits. Among these, at the macro-level, cognitive benefits are the most frequently reported, followed closely by behavioral and meta-cognitive benefits.

Among the specific benefits identified, the most frequently reported one is the behavioral benefit of stimulating revisions to their academic work, ultimately yielding a positive outcome in the form of enhanced writing quality. For instance, Lineback and Holbrook (2023) conducted a rigorous investigation utilizing both qualitative and quantitative methodologies to assess the differences between pre-draft and post-draft versions of students' work. Their analysis encompassed a statistical examination of the scores and an in-depth exploration of students' revision processes and the peer feedback received. The results of this study revealed that 14 out of 15 students experienced an improvement in their overall scores, with 13 students implementing at least one discernible change that could be directly attributed to the influence of peer feedback. Furthermore, in the investigation of the

precise domains exhibiting enhancement subsequent to peer feedback, research has demonstrated that the enhancement of academic writing quality through revision extends to multiple dimensions of academic writing, including but not limited to, the refinement of organizational structure (Kostopoulou and O'Dwyer, 2021; Rodas and Colombo, 2021), the accuracy and appropriateness of in-text citations (Kostopoulou and O'Dwyer, 2021), the depth and clarity of ideas and content (Boillos, 2024; Greenberg, 2015), as well as linguistic precision and appropriateness (Kostopoulou and O'Dwyer, 2021; Zhang et al., 2020).

The second most prevalent advantage, as reported, lies in the meta-cognitive enhancement of self-reflection. Participants generally reported that engaging in peer feedback elicited self-reflection and fostered a more reflective learning approach (Pugh and Veitch, 2019). More precisely, the activity of comparing papers written by different individuals and discussing issues during the peer feedback process prompted students to reflect on their academic writing, with the former activity often functioning spontaneously (Deng et al., 2019). For instance, through the utilization of data sourced from interviews and stimulated recall techniques, Yu's research examining the experiences associated with peer feedback practices during the process of master's thesis writing elucidates that engagement in peer feedback fosters self-reflection upon one's own writing. Through reflection subsequent to critical peer feedback, students strengthened their critical thinking ability and developed into critical readers and writers of academic literature (Ciampa and Wolfe, 2023; Yu, 2019).

Furthermore, some studies have documented that peer feedback activity significantly contribute to the enhancement of students' critical and analytical skills (e.g., Osman et al., 2022), and the confidence in providing constructive peer feedback (Davis, 2014). More specifically, peer feedback practice equips learners with the capacity and confidence to engage in a critical assessment of both their own and their peers' academic work (Boillos, 2024; Davis, 2014; Geithner and Pollastro, 2016; Schillings et al., 2021). For instance, participants in Geithner and Pollastro's (2016) study rated their "ability to provide peer review" significantly higher subsequent to peer feedback practice. Notably, four articles have underscored the superiority of public multi-peer feedback in fostering these essential skills. Specifically, these studies reveal that the diverse perspectives accessible to individual students within the framework of public multi-peer feedback facilitate the identification of overlooked aspects in their own feedback practices, thereby facilitating the refinement and honing of their analytical skills (Gao and Chen, 2024; Chen and Gao, 2024). This underscores the importance of such collaborative feedback mechanisms in nurturing critical thinking and analytical proficiency among students.

It is also noteworthy to highlight the convergence of seven articles, which affirm that the integration of peer feedback into the academic writing process constitutes a significant contributor to the construction of an academic community. Specifically, the interactive exchange during peer feedback sessions, particularly the affective devices embedded in comments, fosters a sense of community among students (Yallop et al., 2021). Furthermore, this practice facilitates the introduction of graduate students into established scholarly networks (Ciampa and Wolfe, 2023; Man et al., 2018; Zhang et al., 2020). For example, Man et al. (2018) examined autonomous peer feedback practices among postgraduate students and observed that such feedback not only catalyzes the construction

of new academic communities but also facilitates the introduction of graduate students into established scholarly networks, echoing the findings of Ciampa and Wolfe (2023) as well as Zhang et al. (2020). In the academic community, peer feedback assumes a pivotal role, serving as a conduit for transmitting academic writing norms and nurturing interpersonal relationships (Zhang et al., 2020). Notably, two recent studies have underscored the distinct advantages of community-based peer feedback in constructing academic community. They emphasized the capacity of this approach to forge social and emotional bonds among classmates, thereby fostering the formation of a cohesive academic community (Gao and Chen, 2024; Chen and Gao, 2024). This underscores the importance of peer feedback not merely as a technical tool but also as a catalyst for building a supportive and collaborative scholarly environment.

Despite many studies elucidating the favorable influence of peer feedback on the revision process, ultimately fostering the advancement of the current writing quality (e.g., Kostopoulou and O'Dwyer, 2021; Rodas and Colombo, 2021), a comparatively scarce body of research has explicitly documented the improvement in writing skills, as evidenced by students' demonstrated capacity to produce high-quality academic writing. Furthermore, within the subset of studies offering such evidence, a substantial proportion relies on self-assessment as the primary metric. For example, López-Pellisa et al. (2021) utilized a 5-point Likert scale to investigate students' perceived enhancement of academic writing proficiency through peer feedback activity. The results of their study indicate that a majority of the participants reported an improvement in their academic writing proficiency subsequent to peer feedback activity. A mere two studies examined the academic writing competence before and after the application of peer feedback (e.g., Hanafi et al., 2024; Ramon-Casas et al., 2019), thereby offering a more objective assessment of the skill enhancement. This paucity of research underscores the need for further investigation to evaluate the impact of peer feedback on the development of academic writing skills.

Despite the multifaceted benefits associated with integrating peer feedback into academic writing, this practice also encounters some challenges, which can be systematically categorized into three distinct categories based on their origins: challenges from peer feedback receivers, challenges from peer feedback providers, and challenges from peer feedback settings.

At a macroscopic level, research has predominantly documented challenges emanating from feedback providers, with subsequent emphasis on those confronted by peer feedback recipients, and finally, challenges inherent in the peer feedback settings. This hierarchical pattern underscores the pivotal role of the two fundamental components of peer feedback activities—the providers and receivers—as the primary sources of issues encountered within this educational practice.

Among the subcategories of challenges, students' deficiency in providing constructive feedback emerges as the paramount obstacle. This underscores a pervasive inability among students to provide insightful peer feedback, a challenge that has been consistently noted across diverse educational contexts, encompassing undergraduate and graduate students at various stages of their academic journey (Álvarez et al., 2015; Gumusoglu et al., 2022). Specifically, feedback was noted to be either insufficient (López-Pellisa et al., 2021; Xu and Zhang, 2023), or characterized by over-generalization, brevity, and superficiality (Cheong et al., 2023; Weaver et al., 2014), neglecting the

intricate issues that truly require attention (Gao et al., 2019). Notably, the quantitative data in Cheong et al.'s (2023) study revealed that 60.8% of the suggestions in peer feedback lacked specificity, minimally contributing to the revision process. This issue can be attributed, in part, to the intricate cognitive and social processing skills required for effective peer feedback generation (Xu and Zhang, 2023), rendering it a formidable task for students to generate constructive feedback on their peers' manuscripts. Furthermore, factors such as limited experience (Yucel et al., 2014), inadequate subject knowledge (Kostopoulou and O'Dwyer, 2021), and constrained metacognitive abilities (Nur and Anas, 2022) have also been identified as contributing factors to this challenge. This deficiency has the potential to engender a lack of trust among students in receiving constructive feedback from their peers (Jurkowski, 2018; Pugh and Veitch, 2019), ultimately impairing their engagement and reducing the efficacy of the peer feedback practice (Álvarez et al., 2015; Jurkowski, 2018).

Consistent with the observed deficiency in delivering constructive feedback, six articles have documented a prevalent lack of confidence among students in providing peer feedback. Notably, Yu's (2019, 2021) research revealed that, despite their enthusiasm for engaging in peer feedback tasks, master's students harbored doubts regarding their linguistic competence, the accuracy and constructiveness of their feedback, and the interpersonal skills necessary for effective peer feedback. Similarly, Ciampa and Wolfe's (2023) study found that doctoral students, despite being advanced academic writers, struggled with perceived inadequacy in their expertise and experience, leading to similar confidence issues. These findings underscore the widespread occurrence of confidence deficits across different academic levels in the context of peer feedback on academic writing. Addressing these confidence deficits is crucial, as they can significantly hinder students' participation in peer feedback, thereby undermining the overall effectiveness of this pedagogical practice (Allen and Katayama, 2016; Xu and Li, 2018; Xue et al., 2023).

Eight articles have consistently highlighted the second most prevalent challenge, which revolves around students' interpersonal apprehensions in offering constructive critiques on their peers' academic writings in non-anonymous settings. Notably, a preponderance of these investigations (specifically, five out of the eight studies) was situated within the Chinese cultural context, where an emphasis on maintaining a harmonious environment is deeply ingrained (Xu and Li, 2018). This psychological pressure can subsequently precipitate a reluctance among students to engage in the peer feedback process (Xu and Li, 2018), or prompt them to grant overly generous grades to their peers (Cheong et al., 2023). For instance, Zhang et al. (2022) reveal that the "face"-threatening dilemmas in the Chinese context cultivate a tendency among students to preserve interpersonal harmony, which often entails an aversion to losing face for their peers. Consequently, it undermines trust among peers and negatively impacts students' willingness to provide constructive feedback, thereby hindering the overall effectiveness of this pedagogical approach.

The deficiency in feedback literacy among receivers also emerges as a significant obstacle in the integration of peer feedback within the domain of academic writing. Feedback literacy, as defined by researchers such as Carless and Boud (2018), encompasses a deep understanding of feedback and effective management, the capacity and disposition to leverage feedback, as well as an appreciation of the

roles of teachers and students themselves in this process. Studies have revealed a tendency among students to selectively incorporate feedback, giving priority to simpler suggestions over more complex ones when revising their academic writing (Shulgina et al., 2024a; Gao et al., 2019; Zhang et al., 2020). For example, in the study of master's students' revision processes, Zhang et al. found a lower revision rate for content-focused feedback (86.11%) compared to form-focused feedback (97.56%). This phenomenon echoes Yu et al.'s (2019) findings, where despite significant behavioral engagement, students lacked strategies and meta-cognitive processing of the feedback, resulting in superficial engagement that hindered the productive use of feedback. These insights emphasize the need for targeted interventions aimed at enhancing students' feedback literacy. By improving their understanding of feedback, empowering them to leverage it effectively, and fostering an appreciation for the peer feedback process, educators can help ensure that the potential of peer feedback is realized in fostering the development of academic writing skills among students.

4.3 Pedagogical implications

The literature under review underscores the multifarious advantages of integrating peer feedback into the process of academic writing. These benefits encompass a broad spectrum, including cognitive enhancements that facilitate academic writing and critical analysis; behavioral improvements marked by active revision in the writing task; affective gains in the form of enhanced self-confidence, heightened willingness, and increased motivation; social benefits stemming from collaborative learning, and a sense of community among peers; as well as meta-cognitive benefits, which are characterized by intensified self-reflection and a heightened meta-cognitive awareness of the writing process, enabling students to better understand and regulate their own writing strategies and approaches. This comprehensive array of benefits underscores the justification for incorporating this method in the development of academic writing.

However, this approach also encounters many challenges, with the most salient being students' deficiency in providing constructive peer feedback, inadequate feedback literacy, and the interpersonal concerns in offering constructive critiques on their peers' academic writings. Drawing upon the insights garnered from the reviewed literature, the subsequent pedagogical interventions are proposed as potential solutions.

Firstly, the provision of comprehensive training on giving peer feedback is paramount to enhancing the overall effectiveness of this practice (Lu et al., 2021; Pugh and Veitch, 2019). Academic writing is an advanced type of writing distinct from conventional school writing, which necessitates a profound grasp of disciplinary knowledge and genre-specific competencies. Students often grapple not just with superficial aspects like vocabulary and grammar but also with advanced facets of academic writing (Gao et al., 2019), including research methodology and the significance of research (Man et al., 2018). However, as evidenced in prior studies, peer feedback tends to focus predominantly on superficial issues, neglecting the more advanced aspects that are central to the purpose of peer feedback in academic writing instruction. Therefore, this training is vital to ensuring that students reap the full benefits of this practice (Chang, 2015; Liou and Peng, 2009).

Extending the discourse, an optimal peer feedback training program for reviewers ought to embody three fundamental elements: the clarification of reviewing criteria (Pugh and Veitch, 2019), the provision of exemplary feedback (Kostopoulou and O'Dwyer, 2021), and the cultivation of a conducive mindset (Yucel et al., 2014). To ensure that reviewers possess a foundational understanding of the pivotal aspects of academic writing, the development of rubrics from the outset is imperative. Rubrics serve as a catalyst for reviewers' engagement in peer feedback (Yu, 2021), enhance their genre-specific knowledge and enable them to generate constructive critiques of their peers' academic work (Ciampa and Wolfe, 2023; Tai et al., 2018). Prior research underscores the positive impact of rubrics in this regard (Ciampa and Wolfe, 2023; Greenberg, 2015; López-Pellisa et al., 2021; Yu, 2021).

A dedicated discussion session focusing on the specific items outlined in the rubrics is recommended. This forum fosters a deeper comprehension of the criteria among students (Yucel et al., 2014), encourages the sharing of insights, and promotes mutual learning (Yu et al., 2019). Additionally, the presentation of exemplary peer feedbacks is vital in illustrating the ideal form of constructive criticism (Costley et al., 2023; Shulgina et al., 2024b). This process should include detailed guidance on feedback-giving strategies, which encompassing emphasizing the importance of addressing advanced issues in academic writing (Gao et al., 2019), prioritizing quality over quantity (Shulgina et al., 2024b), offering comments rather than direct editing (Shulgina et al., 2024a), presenting a diverse range of feedback types that form a coherent logical structure (Lu et al., 2021), and attending to the manner in which feedback is delivered (Lu et al., 2023; Yallop et al., 2021; Zhang et al., 2020).

Furthermore, nurturing a favorable mindset among students is crucial for facilitating the peer feedback process in academic writing context (Yucel et al., 2014). Educators should underscore the potential benefits of both giving and receiving peer feedback, even when students are paired with less proficient peers (Shulgina et al., 2024a). This approach motivates students to engage positively in the activity and helps them establish realistic expectations (Yucel et al., 2014). By addressing these three elements comprehensively, an effective peer feedback training program can be established, thereby maximizing the benefits of this pedagogical practice.

Beyond the refinement of students' feedback skills, an equally pivotal aspect is the cultivation of their feedback literacy, which ultimately determines their ability to reap the full benefits of peer feedback activities (Handley et al., 2011). Therefore, prior to engaging in peer feedback, students must be equipped with strategies to effectively leverage the feedback received (Lu et al., 2021; Yu et al., 2019). This involves teaching them how to incorporate suggestions into their revisions (Álvarez et al., 2015; Shulgina et al., 2024b), fostering feedback acceptance (Lu et al., 2023), and managing diverse types and volumes of feedback effectively (Lu et al., 2023; Shulgina et al., 2024b). Furthermore, students should be guided to participate in the peer feedback process with affective, behavioral, and cognitive engagement to maximize its benefits (Yu et al., 2019).

Although interpersonal concerns frequently emerge in non-anonymous contexts (e.g., Xue et al., 2023; Zhang et al., 2022), the decision to employ anonymity should also go through meticulous consideration. It was found that the anonymity of peer feedback can also deprive opportunities for face-to-face dialogue, which is vital for

elaborating on feedback, fostering constructive commentary, and fostering a sense of responsibility (Schillings et al., 2021). Dialogue not only aids the peer feedback process in clarifying cognitive conflicts (Wu and Lei, 2023), but also in supporting students emotionally (Lineback and Holbrook, 2023), thereby facilitating the revision process. Therefore, alternative methods to mitigate interpersonal tension should be prioritized over anonymous feedback designs. For instance, teachers can impart communication skills that help students manage potentially negative emotions (Zhang et al., 2022).

4.4 Implications on future research

While prior research has provided valuable insights into the benefits and challenges of peer feedback within the realm of academic writing development, its scope is inherently limited in at least four key dimensions.

Firstly, considering the biased attention accorded to undergraduate students in this realm, it is imperative to embark on research endeavors directed toward master's and doctoral students, with the aim of delving into the nuanced potential of peer feedback in academic writing instruction. These studies would not merely elucidate the intricate effects of peer feedback on augmenting postgraduate academic writing skills but also pave the way for a strategic integration of this approach into the development of writing competencies among postgraduate students. Ultimately, such endeavors would contribute significantly to informing the pedagogical implementation of peer feedback in academic writing practices across the entire spectrum of higher education.

Secondly, as highlighted in the preceding section, the quantitative evidence pertaining to the impact of peer feedback remains scarce. Consequently, future research endeavors ought to delve into this topic by conducting rigorous analyses of diverse quantitative datasets, with the aim of providing a more comprehensive and robust understanding of the effects of peer feedback on academic writing development.

Thirdly, while many studies have identified the positive effects of peer feedback on enhancing writing quality, a notable scarcity persists in long-term empirical evidence on the improved writing skills. However, the sustainability and transferability of these effects in fostering students' writing abilities constitute a pivotal aspect in assessing the overall effectiveness of peer feedback (Zhang, 2021). Therefore, future research endeavors ought to employ lagged test to scrutinize the longitudinal effects, thereby elucidating the positive influence of peer feedback on students' academic writing skills.

Lastly, the critical role of sociocultural factors in shaping students' academic development is widely acknowledged, yet previous investigations into peer feedback in academic writing have been notably inadequate in this regard. Out of the reviewed articles, merely five have addressed the influence of cultural factors, and all are confined to the Chinese context, such as the concept of "face" (Zhang et al., 2022) and the tradition of harmonious communication (Xue et al., 2023). Studies that delve into diverse sociocultural backgrounds are expected to contribute significantly to our understanding of this issue, facilitating cross-cultural comparisons and the identification of both similarities and differences in the peer feedback process.

5 Conclusion

The current study employs the PRISMA framework for systematic review to scrutinize the utilization of peer feedback in academic writing. This approach not only maps out the prevailing trends in related research endeavors but also unveils benefits and challenges associated with this practice. The findings of this systematic review reveal a general upward trajectory in research interest, with investigations spanning multiple countries, attesting to the widespread adoption of peer feedback in academic writing pedagogy. However, a notable disparity exists, with a preponderance of studies centering on the formal classroom instruction of undergraduate students' academic writing, as opposed to those focusing on master's and doctoral candidates. Additionally, the preponderance of qualitative data employed in assessing the effects of peer feedback underscores the necessity for future research to adopt a quantitative lens, thereby enriching the understanding of this topic.

The integration of peer feedback in academic writing has been found to yield multifarious benefits, which can be categorized into five distinct domains: cognitive, behavioral, affective, social, and meta-cognitive. Notably, cognitive benefits emerge as the most frequently cited, followed by behavioral benefits, and meta-cognitive benefits. However, the implementation of this approach is not without its challenges, which can be traced to three primary sources: the receiver, the provider, and the setting. Key obstacles encountered include students' inability to provide constructive feedback, a lack of feedback literacy, and interpersonal concerns associated with delivering critical comments. These challenges necessitate careful consideration and strategic interventions to ensure the effective utilization of peer feedback in academic writing instruction.

Despite the systematic review's commendable effort in presenting the prevalent trend and synthesizing the effects of integrating peer feedback into academic writing instruction, thereby offering valuable guidance to both practitioners and researchers in the field, the current study notably confines its focus solely to synthesizing the outcomes of prior investigations. A more profound exploration of the interplay

between the benefits and challenges is anticipated to yield more incisive insights.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

YW: Conceptualization, Investigation, Methodology, Writing – original draft. DL: Conceptualization, Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Ahmed, R. (2021). Peer review in academic writing: different perspectives from instructors and students. *TESOL J.* 12:e00537. doi: 10.1002/tesj.537
- Ahmed, R., and Al-Kadi, A. (2021). Online and face-to-face peer review in academic writing: frequency and preferences. *Eurasian J. Appl. Linguist.* 7, 169–201. doi: 10.32601/ejal.911245
- Aitchison, C., and Lee, A. (2006). Research writing: problems and pedagogies. *Teach. High. Educ.* 11, 265–278. doi: 10.1080/13562510600680574
- Allen, D., and Katayama, A. (2016). Relative second language proficiency and the giving and receiving of written peer feedback. *System.* 56, 96–106. doi: 10.1016/j.system.2015.12.002
- Álvarez, M., Boillos, M. M., and Velasco, E. (2015). The process of building the abstract of a paper in the context of a writing workshop for PhDs. *Procedia. Soc. Behav. Sci.* 178, 2–8. doi: 10.1016/j.sbspro.2015.03.137
- Andrews, R., and Harlen, W. (2006). Issues in synthesizing research in education. *Educ. Res.* 48, 287–299. doi: 10.1080/00131880600992330
- Aromataris, E., and Pearson, A. (2014). The systematic review: an overview. *Am. J. Nurs.* 114, 53–58. doi: 10.1097/01.NAJ.0000444496.24228.2c
- Bearman, M., Smith, C. D., Carbone, A., Slade, S., Baik, C., Hughes-Warrington, M., et al. (2012). Systematic review methodology in higher education. *High. Educ. Res. Dev.* 31, 625–640. doi: 10.1080/07294360.2012.702735
- Boillos, M. M. (2024). Peer review in early academic writing: impact vs. students' beliefs. *Innov. Lang. Learn. Teach.* 18, 402–410. doi: 10.1080/17501229.2024.2311836
- Bruffee, K. A. (1984). Collaborative learning and the "conversation of mankind". *Coll. Engl.* 46, 635–652. doi: 10.2307/376924
- Carless, D., and Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assess. Eval. High. Educ.* 43, 1315–1325. doi: 10.1080/02602938.2018.1463354
- Chakraborty, D., Soyoof, A., Moharami, M., Utami, A. D., Zeng, S., Cong-Lem, N., et al. (2021). Feedback as a space for academic social practice in doctoral writing groups. *Educ. Dev. Psychol.* 38, 238–248. doi: 10.1080/20590776.2021.1972764
- Chalmers, I., Hedges, L. V., and Cooper, H. (2002). A brief history of research synthesis. *Eval. Health Prof.* 25, 12–37. doi: 10.1177/0163278702025001003
- Chang, C. Y. (2015). Teacher modeling on EFL reviewers' audience-aware feedback and affectivity in L2 peer review. *Assess. Writ.* 25, 2–21. doi: 10.1016/j.ASW.2015.04.001
- Chen, W., and Gao, J. (2024). Creating an online community of inquiry: learner practice and perceptions of community-based feedback giving in academic writing. *Comput. Assist. Lang. Learn.* 37, 493–520. doi: 10.1080/09588221.2022.2043381
- Cheong, C. M., Luo, N., Zhu, X., Lu, Q., and Wei, W. (2023). Self-assessment complements peer assessment for undergraduate students in an academic writing task. *Assess. Eval. High. Educ.* 48, 135–148. doi: 10.1080/02602938.2022.2069225

- Ciampa, K., and Wolfe, Z. M. (2023). From isolation to collaboration: creating an intentional community of practice within the doctoral dissertation proposal writing process. *Teach. High. Educ.* 28, 487–503. doi: 10.1080/13562517.2020.1822313
- Colombo, L., and Rodas, E. L. (2021). Interdisciplinarity as an opportunity in Argentinian and Ecuadorian writing groups. *High. Educ. Res. Dev.* 40, 207–219. doi: 10.1080/07294360.2020.1756750
- Costley, J., Zhang, H., Courtney, M., Shulgina, G., Baldwin, M., and Fanguy, M. (2023). Peer editing using shared online documents: the effects of comments and track changes on student L2 academic writing quality. *Comput. Assist. Lang. Learn.* 1–27, 1–27. doi: 10.1080/09588221.2023.2233573
- Davies, P. (2000). The relevance of systematic reviews to educational policy and practice. *Oxf. Rev. Educ.* 26, 365–378. doi: 10.1080/713688543
- Davis, L. E. (2014). A workshop series using peer-grading to build drug information, writing, critical-thinking, and constructive feedback skills. *Am. J. Pharm. Educ.* 78:191. doi: 10.5688/ajpe7810191
- Deng, Y., Kelly, G. J., and Deng, S. (2019). The influences of integrating reading, peer evaluation, and discussion on undergraduate students' scientific writing. *Int. J. Sci. Educ.* 41, 1408–1433. doi: 10.1080/09500693.2019.1610811
- Eppler, E., Meyer, J., Serowy, S., Link, K., Pauk, B., and Filgueira, L. (2021). Enhancing scientific communication skills: A real-world simulation in a tertiary-level life science class using E-learning technology in biomedical literature perception, reflective review writing on a clinical issue, and self and peer assessments. *Res. Sci. Educ.* 51, 277–299. doi: 10.1007/s11165-018-9795-7
- Gao, J., and Chen, W. (2024). Developing culturally-situated student feedback literacy through multi-peer feedback giving: an online community-based approach. *Lang. Aware.* 2024, 1–24. doi: 10.1080/09658416.2024.2321894
- Gao, Y., Schunn, C. D. D., and Yu, Q. (2019). The alignment of written peer feedback with draft problems and its impact on revision in peer assessment. *Assess. Eval. High. Educ.* 44, 294–308. doi: 10.1080/02602938.2018.1499075
- Geithner, C. A., and Pollastro, A. N. (2016). Doing peer review and receiving feedback: impact on scientific literacy and writing skills. *Adv. Physiol. Educ.* 40, 38–46. doi: 10.1152/advan.00071.2015
- Goh, C. F., Tan, O. K., Rasli, A., and Choi, S. L. (2019). Engagement in peer review, learner-content interaction and learning outcomes. *Int. J. Inf. Learn. Technol.* 36, 423–433. doi: 10.1108/IJILT-04-2018-0038
- Gondim, S. M., and Mutti, C. (2011). Affections in learning situations: a study of an entrepreneurship skills development course. *J. Work. Learn.* 23, 195–208. doi: 10.1108/13665621111117224
- Goupil, L., and Kouider, S. (2019). Developing a reflective mind: from core metacognition to explicit self-reflection. *Curr. Dir. Psychol. Sci.* 28, 403–408. doi: 10.1177/0963721419848672
- Greenberg, K. P. (2015). Rubric use in formative assessment: A detailed behavioral rubric helps students improve their scientific writing skills. *Teach. Psychol.* 42, 211–217. doi: 10.1177/0098628315587618
- Gumusoglu, S. B., Soulinthavong, M. N., and Barr, J. (2022). A WINning approach: teaching science communication skills through small-group workshops. *High. Learn. Res. Commun.* 12, 23–44. doi: 10.18870/hlrc.v12i2.1342
- Hanafi, I., Kheder, K., Sabouni, R., Gorra al Nafouri, M., Hanafi, B., Alsalkini, M., et al. (2024). Improving academic writing in a low-resource country: a systematic examination of online peer-run training. *Teach. Learn. Med.* 29, 1–15. doi: 10.1080/10401334.2024.2332890
- Handley, K., Price, M., and Millar, J. (2011). Beyond 'doing time': investigating the concept of student engagement with feedback. *Oxf. Rev. Educ.* 37, 543–560. doi: 10.1080/03054985.2011.604951
- Hsieh, H.-F., and Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qual. Health Res.* 15, 1277–1288. doi: 10.1177/1049732305276687
- Huisman, B., Saab, N., Van Den Broek, P., and Van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: A meta-analysis. *Assess. Eval. High. Educ.* 44, 863–880. doi: 10.1080/02602938.2018.1545896
- Hyland, K. (2004). *Disciplinary discourses: social interactions in academic writing*. Ann Arbor, MI: University of Michigan Press.
- Jurkowski, S. (2018). Do question prompts support students in working with peer feedback? *Int. J. Educ. Res.* 92, 1–9. doi: 10.1016/j.ijer.2018.07.003
- Kabaran, G. G. (2022). Graduate students' perceptions of the academic writing process and its problems: a qualitative study in Turkey. *Issues Educ. Res.* 32, 943–959.
- Kondracki, N. L., and Wellman, N. S. (2002). Content analysis: review of methods and their applications in nutrition education. *J. Nutr. Educ. Behav.* 34, 224–230. doi: 10.1016/s1499-4046(06)60097-3
- Kostopoulou, S., and O'Dwyer, F. (2021). "We learn from each other": peer review writing practices in English for academic purposes. *Lang. Learn. High. Educ.* 11, 67–91. doi: 10.1515/cercles-2021-2006
- Lineback, J. E., and Holbrook, E. (2023). Engaging in a collaborative space: exploring the substance and impact of peer review conversations. *Assess. Eval. High. Educ.* 49, 556–571. doi: 10.1080/02602938.2023.2290978
- Liou, H. C., and Peng, Z. Y. (2009). Training effects on computer-mediated peer review. *System.* 37, 514–525. doi: 10.1016/j.system.2009.01.005
- Liu, F., Du, J., Zhou, D. Q., and Huang, B. (2021). Exploiting the potential of peer feedback: the combined use of face-to-face feedback and e-feedback in doctoral writing groups. *Assess. Writ.* 47:100482. doi: 10.1016/j.asw.2020.100482
- Liu, N. F., and Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teach. High. Educ.* 11, 279–290. doi: 10.1080/13562510600680582
- Liu, S., Liu, S., Liu, Z., Peng, X., and Yang, Z. (2022). Automated detection of emotional and cognitive engagement in MOOC discussions to predict learning achievement. *Comp. Educ.* 181:104461. doi: 10.1016/j.compedu.2022.104461
- López-Pellisa, T., Rotger, N., and Rodríguez-Gallego, F. (2021). Collaborative writing at work: peer feedback in a blended learning environment. *Educ. Inf. Technol.* 26, 1293–1310. doi: 10.1007/s10639-020-10312-2
- Lu, Q., Yao, Y., and Zhu, X. (2023). The relationship between peer feedback features and revision sources mediated by feedback acceptance: the effect on undergraduate students' writing performance. *Assess. Writ.* 56:100725. doi: 10.1016/j.asw.2023.100725
- Lu, Q., Zhu, X., and Cheong, C. M. (2021). Understanding the difference between self-feedback and peer feedback: a comparative study of their effects on undergraduate students' writing improvement. *Front. Psychol.* 12:739962. doi: 10.3389/fpsyg.2021.739962
- Man, D., Xu, Y., and O'Toole, J. M. (2018). Understanding autonomous peer feedback practices among postgraduate students: A case study in a Chinese university. *Assess. Eval. High. Educ.* 43, 527–536. doi: 10.1080/02602938.2017.1376310
- Martin, F., Dennen, V. P., and Bonk, C. J. (2020). A synthesis of systematic review research on emerging learning environments and technologies. *Educ. Technol. Res. Dev.* 68, 1613–1633. doi: 10.1007/s11423-020-09812-2
- Moses, L. J., and Baird, J. A. (1999). "Metacognition" in *The MIT encyclopaedia of the cognitive sciences*. eds. R. A. Wilson and F. C. Keil (Cambridge, MA: MIT), 533–535.
- Nazamud-din, A., Hariz Zaini, M., and Mohd Jamil, N. H. (2020). The relationship of affective, behavioral and cognitive engagements in ESL higher learning classroom. *Eng. Lang. Teach. Linguist. Stud.* 2, p48–p64. doi: 10.22158/eltls.v2n4p48
- Nur, S., and Anas, I. (2022). Engaging novice writers in online collaborative review through peer-review circles. *Int. J. Lang. Educ.* 6, 63–74. doi: 10.26858/ijole.v6i1.26141
- Opfer, V., and Pedder, D. (2011). Conceptualizing teacher professional learning. *Rev. Educ. Res.* 81, 376–407. doi: 10.3102/0034654311413609
- Osman, A., Al-Badriyeh, D., Hussain, F. N., Riaz, S., Elewa, H., and Mraiche, F. (2022). The design and implementation of an undergraduate health professional degree elective course on scientific writing, peer assessment, and critical appraisal. *Curr. Pharm. Teach. Learn.* 14, 765–772. doi: 10.1016/j.cptl.2022.06.008
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71. doi: 10.1136/bmj.n71
- Pearson, A. (2004). Balancing the evidence: incorporating the synthesis of qualitative data into systematic reviews. *JBI Reports* 2, 45–64. doi: 10.1111/j.1479-6988.2004.00008.X
- Piaget, J. (1962). The relation of affectivity to intelligence in the mental development of the child. *Bull. Menn. Clin.* 26, 129–137.
- Potvin, S., Pelletier, J., Grot, S., Hébert, C., Barr, A. M., and Lecomte, T. (2018). Cognitive deficits in individuals with methamphetamine use disorder: a meta-analysis. *Addict. Behav.* 80, 154–160. doi: 10.1016/j.addbeh.2018.01.021
- Pugh, M., and Veitch, F. (2019). Undergraduate peer review, reading and writing: reflecting on experiences from an international politics module. *Eur. Political Sci.* 18, 335–350. doi: 10.1057/s41304-018-0178-3
- Ramon-Casas, M., Nuño, N., Pons, F., and Cunillera, T. (2019). The different impact of a structured peer-assessment task in relation to university undergraduates' initial writing skills. *Assess. Eval. High. Educ.* 44, 653–663. doi: 10.1080/02602938.2018.1525337
- Rodas, E. L., and Colombo, L. (2021). Self-managed peer writing groups for the development of EFL literacy practices. *TESL-EJ* 24, 1–22.
- Sallee, M. W., Hallett, R. E., and Tierney, W. G. (2011). Teaching writing in graduate school. *Coll. Teach.* 59, 66–72. doi: 10.1080/87567555.2010.511315
- Santelmann, L. M., Stevens, D. D., and Martin, S. B. (2018). Fostering master's students' metacognition and self-regulation practices for research writing. *Coll. Teach.* 66, 111–123. doi: 10.1080/87567555.2018.1446898
- Schillings, M., Roebertsen, H., Savelberg, H., and Dolmans, D. (2023). A review of educational dialogue strategies to improve academic writing skills. *Act. Learn. High. Educ.* 24, 95–108. doi: 10.1177/1469787418810663
- Schillings, M., Roebertsen, H., Savelberg, H., van Dijk, A., and Dolmans, D. (2021). Improving the understanding of written peer feedback through face-to-face peer dialogue: students' perspective. *High. Educ. Res. Dev.* 40, 1100–1116. doi: 10.1080/07294360.2020.1798889
- Shulgina, G., Costley, J., Shcheglova, I., Zhang, H., and Sedova, N. (2024a). Online peer editing: the influence of comments, tracked changes and perception of participation on students' writing performance. *Smart Learn. Environ.* 11:30. doi: 10.1186/s40561-024-00315-8

- Shulgina, G., Fanguy, M., Zhang, H., Courtney, M., Baldwin, M., and Costley, J. (2024b). The moderating effects of total comments on the relationship between comment implementation and online peer-supported writing performance. *Comp. Educ.* 219:105104. doi: 10.1016/j.compedu.2024.105104
- Siddaway, A. P., Wood, A. M., and Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annu. Rev. Psychol.* 70, 747–770. doi: 10.1146/annurev-psych-010418-102803
- Singleton-Jackson, J., and Lumsden, D. B. (2009). Johnny still can't write, even if He Goes to college: a study of writing proficiency in higher education graduate students. *Curr. Issues Educ.* 121, 1–39.
- Svalberg, A. M.-L. (2009). Engagement with language: interrogating a construct. *Lang. Aware.* 18, 242–258. doi: 10.1080/09658410903197264
- Swain, M. (2013). The inseparability of cognition and emotion in second language learning. *Lang. Teach.* 46, 195–207. doi: 10.1017/S0261444811000486
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge, UK: Cambridge University Press.
- Swales, J. M., and Feak, C. B. (1994). *Academic writing for graduate students: Essential tasks and skills*. Ann Arbor, MI: University of Michigan Press.
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., and Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *High. Educ.* 76, 467–481. doi: 10.1007/s10734-017-0220-3
- Topping, K. J. (2009). Peer Assessment. *Theory Pract.* 48, 20–27. doi: 10.1080/00405840802577569
- Uher, J. (2016). What is behaviour? And (when) is language behaviour? A metatheoretical definition. *J. Theory Soc. Behav.* 46, 475–501. doi: 10.1111/jtsb.12104
- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Weaver, C., Duran, E., and Nikles, J. (2014). An integrated approach for development of scientific writing skills in under-graduate organic lab. Addressing the millennial student in undergraduate chemistry. *ACS Symp. Ser.* 1180, 105–123. doi: 10.1021/bk-2014-1180.ch008
- Wu, P., and Lei, C. (2023). Developing feedback literacy through dialogue-supported performances of multi-draft writing in a postgraduate class. *Assess. Writ.* 57:100759. doi: 10.1016/j.asw.2023.100759
- Xue, S., Yang, Y., Du, J., and Liu, F. (2023). Multi-layered e-feedback anxiety: an action research study among Chinese learners using peer feedback activities in an academic writing course. *Front. Psychol.* 14:1062517. doi: 10.3389/fpsyg.2023.1062517
- Xu, L., and Zhang, T. (2023). Engaging with multiple sources of feedback in academic writing: postgraduate students' perspectives. *Assess. Eval. High. Educ.* 48, 995–1008. doi: 10.1080/02602938.2022.2161089
- Xu, X., and Li, X. (2018). Teaching academic writing through a process-genre approach: a pedagogical exploration of an EAP program in China. *TESL-EJ* 22:n2.
- Yallop, R. M. A., Taremaa, P., and Leijen, D. A. J. (2021). The affect and effect of asynchronous written feedback comments on the peer feedback process: an ethnographic case-study approach within one L2 English doctorate writing group. *J. Writ. Res.* 12, 531–600. doi: 10.17239/jowr-2021.12.03.02
- Yucel, R., Bird, F. L., Young, J., and Blanksby, T. (2014). The road to self-assessment: exemplar marking before peer review develops first-year students' capacity to judge the quality of a scientific report. *Assess. Eval. High. Educ.* 39, 971–986. doi: 10.1080/02602938.2014.880400
- Yu, S. (2019). Learning from giving peer feedback on postgraduate theses: voices from master's students in the Macau EFL context. *Assess. Writ.* 40, 42–52. doi: 10.1016/j.asw.2019.03.004
- Yu, S. (2021). Giving genre-based peer feedback in academic writing: sources of knowledge and skills, difficulties and challenges. *Assess. Eval. High. Educ.* 46, 36–53. doi: 10.1080/02602938.2020.1742872
- Yu, S., and Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Lang. Teach.* 49, 461–493. doi: 10.1017/S0261444816000161
- Yu, S., Zhang, Y., Zheng, Y., Yuan, K., and Zhang, L. (2019). Understanding student engagement with peer feedback on master's theses: a Macau study. *Assess. Eval. High. Educ.* 44, 50–65. doi: 10.1080/02602938.2018.1467879
- Zhang, M., He, Q., Du, J., Liu, F., and Huang, B. (2022). Learners' perceived advantages and social-affective dispositions toward online peer feedback in academic writing. *Front. Psychol.* 13:973478. doi: 10.3389/fpsyg.2022.973478
- Zhang, T. (2021). The effect of highly focused versus mid-focused written corrective feedback on EFL learners' explicit and implicit knowledge development. *System* 99:102493. doi: 10.1016/j.system.2021.102493
- Zhang, Y., Yu, S., and Yuan, K. (2020). Understanding master's students' peer feedback practices from the academic discourse community perspective: a rethinking of postgraduate pedagogies. *Teach. High. Educ.* 25, 126–140. doi: 10.1080/13562517.2018.1543261
- Zheng, L., Zhang, X., and Cui, P. (2019). The role of technology-facilitated peer assessment and supporting strategies: a meta-analysis. *Assess. Eval. High. Educ.* 45, 372–386. doi: 10.1080/02602938.2019.1644603



OPEN ACCESS

EDITED BY

Xinghua Liu,
Shanghai Jiao Tong University, China

REVIEWED BY

Shizhou Yang,
Payap University, Thailand
Ruoxi Liu,
Peking University, China
Ye Yuan,
University of Science Malaysia (USM), Malaysia
Yaping Guo,
Beihang University, China

*CORRESPONDENCE

Donghong Liu
✉ liudonghong19@hotmail.com

RECEIVED 22 September 2024

ACCEPTED 27 November 2024

PUBLISHED 10 December 2024

CITATION

Tian Y and Liu D (2024) A bibliometric study of identity construction in English writing for academic purposes.

Front. Psychol. 15:1499917.

doi: 10.3389/fpsyg.2024.1499917

COPYRIGHT

© 2024 Tian and Liu. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A bibliometric study of identity construction in English writing for academic purposes

Yuhan Tian and Donghong Liu*

School of Foreign Languages, Southeast University, Nanjing, China

Identity construction is a crucial factor in assessing and enhancing the quality of academic writing. However, identity is elusive and difficult to capture due to its abstract nature. Most existing literature discussed academic writing in a general way, overlooking specific studies on identity construction in articles, theses, and dissertations. This study aims to provide a comprehensive review of studies on identity construction in typical academic writing and assist readers in understanding the development, discoveries, and future trends in this field. It seeks to enlighten scholars and students regarding future research directions and to improve academic writing quality in practice. A bibliometric tool, *CiteSpace*, was used together with manual close reading. The data were primarily retrieved from the Web of Science database. Keyword co-occurrence and cluster analyses were conducted to describe the current state of research and predict future hotspots. It was found that the literature in this field generally showed an upward trend before 2020. High-frequency keywords primarily relate to literacy, doctoral education, pedagogy, plagiarism, and gender, representing this field's primary research area. Most clusters exhibit a high level of novelty but have not yet received the attention they deserve because they are situated in the second quadrant of the coordinate diagram as potential clusters. Clusters focusing on socio-cultural identity and the pedagogy of identity construction are more prominent than the other areas. Those focusing on academic (professional) development related to authorial and academic identity are more novel.

KEYWORDS

bibliometric study, identity construction, English writing, academic purpose, *CiteSpace*

1 Introduction

Identity has long been a prominent topic in academic writing within higher education. Ivanič (1998) categorized the various textual identities of a writer that require integration, including the autobiographical self, discursive self, self as author, and possibilities for self-hood. More specifically, identity can be categorized into three dimensions: authorial identity constructed in text, academic (professional) identity developed in higher education, and socio-cultural identity reflected in the writing. The authorial identity is defined as “the sense a writer has of themselves as an author and the textual identity they construct in their writing” (Pittam et al., 2009, p. 154). Academic (professional) identity refers to the researchers' recognition of themselves as a part of the academic community (Botelho de Magalhães et al., 2019). Authors' socio-cultural identities encompass their authentic individual identities within society, such as gender, class, and ethnicity, which inevitably influence and are reflected in their academic identity and writing.

Authorial identity contributes to a credible and professional image for authors, thereby enhancing the persuasiveness and quality of their articles. Significant progress has been made

in studies on authorial identity, primarily as a pedagogical approach to help students improve academic writing in higher education. It has been demonstrated that enhancing students' authorial identity could help prevent plagiarism in academic writing (Khoo and Kang, 2022). Various measures of authorial identity have been discussed to ensure the effective prevention of plagiarism and shed light on the practical implications of authorial identity construction (Cheung et al., 2017; Ballantine et al., 2018).

Academic (professional) identity determines the extent to which writers are willing to conform to the norms and conventions of academic writing (Botelho de Magalhães et al., 2019). Studies on academic identity primarily adopt a scholarly approach, focusing on the academic development of scholars in higher education. Academic identity is continuously constructed and reconstructed across various higher education institutions as part of professional development (French, 2020). This spatial dynamics of academic identity is shaped by the spatial nature of academic writing practices (Beighton, 2020), embodied in writing groups and retreats. These practices greatly benefit participants' academic development and well-being, involving doctoral students, academic returnees, novice lecturers, and experienced teachers.

A comprehensive and nuanced understanding of socio-cultural identities leads to authors' more mature and flexible use of the language reservoir in their academic writing. Studies on authors' socio-cultural identities center on the issue of how aspects such as language, gender, class, and ethnicity are reflected in and influence their academic identity and writing. The language aspect in the context of multilingualism is particularly significant due to the globalization of the academic community. Authors' multilingual skills should be seen as valuable resources rather than challenges to overcome in the process of constructing their identities, which can effectively prevent feelings of inferiority and inequality (Liu and Tannacito, 2013).

The solid construction and skillful integration of identities across all three dimensions mentioned above in academic writing contribute to producing high-quality articles and theses. Writing enables greater deliberation and precision than speaking, allowing identity to be effectively embodied and captured. It serves as the primary means of expressing viewpoints and engaging with audiences in academic contexts. Identity construction involves both the approach and the purpose of enhancing the quality and persuasiveness of academic writing.

Identity construction is a significant factor influencing reviewers' evaluations of writing quality. High-quality academic output is crucial for scholars' professional development and students' attainment of degrees in higher education. Higher education institutions require scholars to publish high-quality articles in reputable journals. Similarly, students must continuously improve their academic writing and complete a demanding thesis or dissertation to attain their degrees.

Raitskaya and Tikhonova (2022) have pointed out that "identity" together with "teaching and learning academic writing in higher education," "writing for publication," and "writing a thesis" are among the 25 most frequent keywords in their systematic review of academic writing. However, no systematic review has been conducted on the combined field of these four keywords, specifically focusing on identity construction in articles, theses, and dissertations writing in higher education. Academic writing, including publications and theses, is essential for scholars and students in higher education. The

effective construction and strategic use of identity significantly enhance the quality of academic writing, thereby fostering their development within the academic environment. And most of the existing literature employed qualitative methods, such as narrative analysis and thematic analysis, to examine data from comparative cases, questionnaires, and interviews, while also using autobiography and duoethnography to reflect individual perspectives and experiences. Therefore, their research findings carry a certain degree of subjectivity. The quantitative characteristics of this bibliometric study enhance its objectivity. Given its practical and theoretical implications, reviewing the combined field is necessary and significant.

In light of this gap, this study will provide a comprehensive summary and analysis of current research findings using the bibliometric tool *CiteSpace*, offering a more objective understanding of the development, discoveries, and future trends in identity construction in English academic writing in higher education. The focus will be on traditional academic writing, including articles, theses, and dissertations, which exhibit the most salient characteristics of scholarly work. Atypical genres, such as reflective writing and autoethnography, will only be deemed approaches to explore typical academic production in the screened articles.

Specifically, this study aims to address the following three questions:

- (1) What are the stages of development in the studies on identity construction in academic writing, and what is the general trend?
- (2) What are the focal areas in the studies on identity construction?
- (3) What are the current hotspots and potential future ones?

The findings of this study contribute not only to a deeper understanding of identity construction in English academic writing but also to the effective enhancement of writing quality and development for scholars and students, holding great theoretical and practical significance.

2 Methods

2.1 Data sources, search strategy, and data extraction

First, English-language journal articles from the Core Collections of Web of Science (WoS) database were retrieved as data from *CiteSpace* and close reading. Web of Science is a comprehensive research database renowned for its rigorous indexing and high-quality content. It provides valuable resources for researchers through its advanced search features and extensive coverage. The Core Collection within Web of Science comprises a curated selection of high-quality journals that feature reliable and impactful studies. To ensure the completeness of data and the integrity of the publication year, the time span was set from 1992.1.1, when articles became available for retrieval in this database, to 2023.12.31. The terms "identity," "academic writing," and "higher education" were selected to retrieve relevant articles. Additionally, we limited the retrieved data to the type "article" and the language "English" to ensure that only English articles were obtained. Second, citation searches from the references of the above-acquired articles and manual searches from Google Scholar were also

employed to supplement any omitted literature. Five more articles have been searched and added to the entire dataset, as shown in [Figure 1](#).

Manual close reading was interspersed throughout the entire study process to screen the targeted literature according to inclusion and exclusion criteria and to generalize conclusions to supplement the results obtained from running *CiteSpace* more elaborately.

To accurately target the articles for review, the retained articles should meet the requirements that (1) studies were related to articles, theses, and dissertation writing; (2) studies were targeted to identity construction; (3) studies were involved in the environment of higher education. Out of 192 articles retrieved from WoS database, 131 irrelevant documents were manually eliminated, owing to their irrelevant foci, including (1) studies without focus on articles, theses, and dissertation writing but on some other atypical academic genres such as reflective writing; (2) studies only focusing on construction of professional identities such as lawyer and tutor or autobiographical identities such as gender, ethnic, and racial identities but failing to relate identity with academic writing; (3) studies without focus on higher education level.

Both authors conducted the screening processes independently. After the respective screenings, we compared the results, identified any inconsistencies, and reached an agreement through negotiation to determine the final articles for analysis. In the end, 61 pieces of data from the WoS database were retained for the final analysis. Five additional articles from citation references and Google searches were included, totaling 66 articles considered in the final analysis. The detailed information for all 66 articles is presented in [Supplementary material](#).

2.2 Procedure

First, the data from WoS underwent keyword co-occurrence and cluster analysis by *CiteSpace*. It stands out as a robust bibliometrics tool employed extensively for data analysis and visualization. It can handle bibliographic and citation information sourced from major databases like the Web of Science ([Chen, 2006](#)).

A knowledge map will be produced through co-occurrence analysis to illustrate the knowledge structure regarding high-frequency keywords. This will enable us to gain a deeper and more comprehensive understanding of this field's current state of research.

The relevant literature will be divided into clusters through cluster analysis, revealing their internal structure. These categorized clusters will be plotted on a strategic coordinate graph, with popularity and novelty represented on the horizontal and vertical axes. This will clearly present current research hotspots (clusters in the first quadrant) and predict potential future hotspots (clusters in the second quadrant).

Second, data obtained through citation searches from the references of previously acquired articles and manual searches on Google Scholar underwent close reading to supplement any omissions found in the literature acquired from WoS. Close reading is integrated throughout the entire research process in conjunction with the *CiteSpace* tool.

3 Results

3.1 Annual distribution

All these articles were distributed over 21 years, as shown in [Figure 2](#), starting from 2003, when the first article was published.

It can be concluded that before 2016, articles in this field showed a slight increase amidst fluctuations. The overall data was relatively small, with only 18 articles produced over the 14 years. From 2017 to 2020, the number of articles increased significantly compared to the previous stage. It sharply increased starting in 2017 and peaked in 2020. Since then, it experienced a sharp decline until 2022. From the dotted trend line, it can be seen that the literature in this field generally showed an upward trend before 2020.

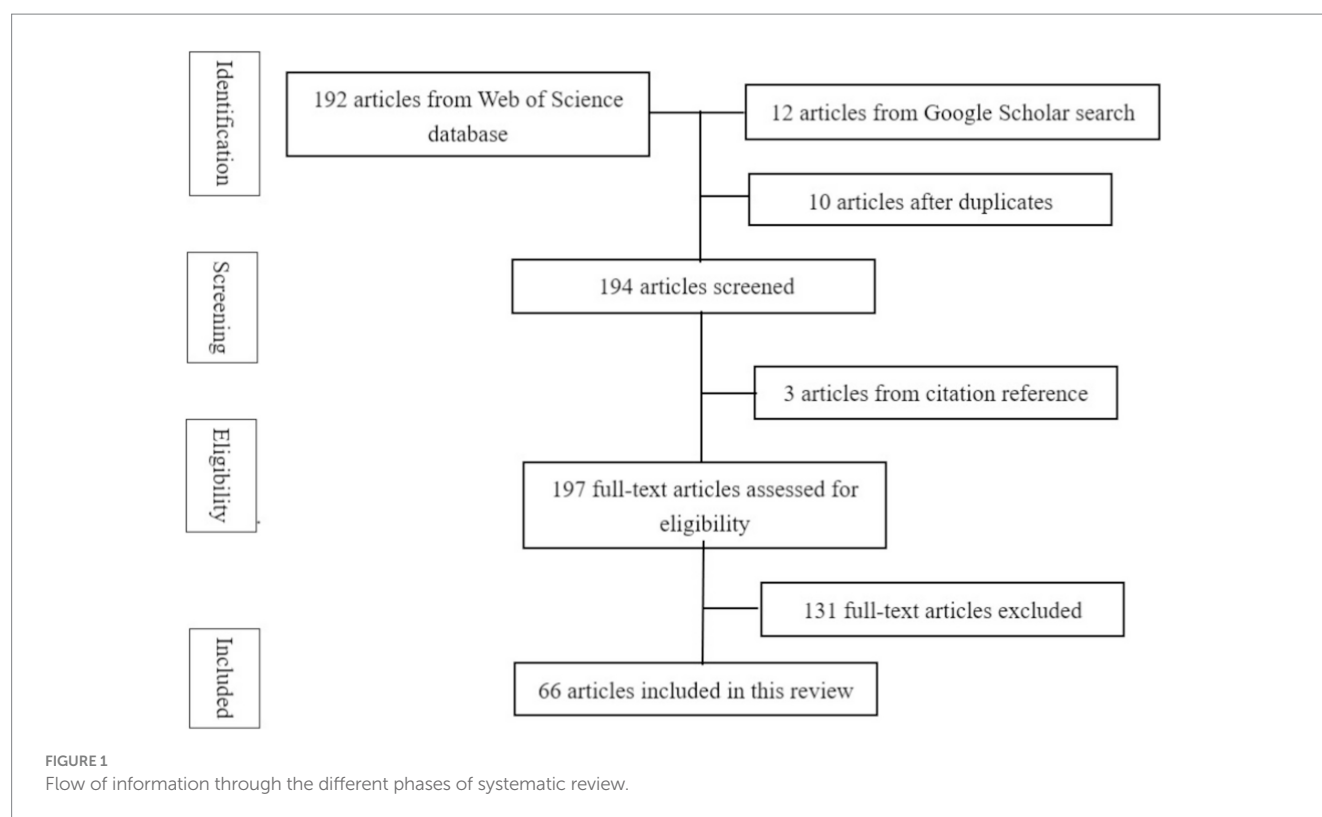
3.2 Visualization of keyword co-occurrence

CiteSpace was employed as the tool to analyze data from WoS. The selection criteria were set by thresholds, adjusted multiple times to (1, 1, 10) (1, 1, 10) (1, 1, 10). Two hundred and seventy two nodes and 1,102 lines were obtained, which yielded a line-to-node ratio of 1.5 and provided a solid and appropriate basis for analysis. These nodes and lines were depicted in [Figure 3](#) and visualized over a timeline in [Figure 4](#). The larger font of the node and character in [Figure 3](#) indicated a higher frequency of keywords. A keyword burst detection analysis was conducted, and the results are presented in [Figure 5](#), illustrating a surge in keyword frequency. [Table 1](#) displays all keywords with a frequency of three or more and the initial year of their appearance, excluding search terms.

3.2.1 Studies focusing on academic writing

It can be concluded from [Table 1](#) that in terms of academic writing, studies are primarily involved in the subfield of literacy concerning high-frequency keywords such as “academic literacy,” “literacy,” and “multiliteracy.” They are closely related to the increasingly diverse population in higher education. According to [Figure 5](#), both “academic literacy” and “literacy” show burst strength, with “academic literacy” exhibiting significantly higher strength than the other keywords and lasting for an extended period. This underscores its importance in the development of this field. Although “multiliteracy” shows no burst strength in [Figure 5](#), it appears relatively early in [Figure 4](#), demonstrating its fundamental role in this field.

In academic writing, literacy is the skillful use of language to understand and communicate a writer's viewpoints effectively. As a perspective of academic writing, academic literacy ([Lea and Street, 2006](#)) is proposed as an alternative to the deficit model and the academic socialization model in response to more flexible and accessible higher education institutions with increasingly diverse student populations. This approach moves beyond traditional views of writing as a set of discrete technical skills in need of fixing, such as grammar or spelling, and instead considers writing as a social practice that varies according to context, culture, and genre. It is related to social identity, power, authority, and meaning-making ([Lea and Street, 2006](#)), which emphasizes that students' academic writing is deeply connected to their understanding of knowledge and the development of their identities. Besides, academic writing is viewed as dynamic and contested, with diverse interpretations of “good writing” across contexts, and critical awareness of the hidden assumptions and values underlying academic practices is encouraged. Academic literacy theory ([Lea and Street, 2006](#)) provides a comprehensive framework



for understanding the complexities of academic writing in higher education.

Multiliteracy acknowledges the diverse ways people communicate and make meaning in the increasingly globalized world, emphasizing the importance of understanding and valuing diverse languages, dialects, and cultural practices in communication. Multilingual and translingual writers are significant subjects in this field (Canagarajah, 2020). They are typically reflective, knowledgeable, and skilled individuals with transnational or transethnic experiences, adept at navigating the diverse language and literacy environments they encounter in their daily university lives. However, they often face challenges in adapting their multilingual and multiliteracy skills to fit the academic context, occasionally highlighting or downplaying these abilities as needed. Canagarajah (2015) examined how a dialogical pedagogy in multilingual writing classrooms helped students construct their academic voices through a case study. His study further demonstrated that academic writing voice was a hybrid construct that involved negotiating personal identity, cultural background, and academic conventions.

3.2.2 Studies focusing on higher education

Sub-topics focusing on higher education are the second largest part of the literature according to high-frequency keywords in Table 1, including “doctoral student,” “doctoral education,” “university,” “pedagogy,” and “student.” Based on Figures 4, 5, keywords in this field emerged early, accompanied by various bursts of keyword activity throughout its development, such as “writing group” and “peer feedback” in recent years, implying that research in this general field is relatively vibrant.

Doctoral education in universities is the primary research focus, although some studies examine undergraduates as research subjects.

Academic literacy provides a frame for designing curriculum and instruction, emphasizing pedagogy in universities (Lea and Street, 2006). Exploring effective pedagogical approaches to improving academic writing is another crucial area of interest within this field. Social support mechanisms are particularly significant in the academic literacy model because they directly address the challenges of navigating writing as a social practice. For instance, writing centers, as a form of writing support, are among the most significant resources. An academic writing center can serve as a mentoring environment and a collaborative learning space, helping students and young academics explore their academic identities and facilitating their transition and transformation (Archer and Parker, 2016).

3.2.3 Studies focusing on identity

Other high-frequency keywords in Table 1 are related to identity concerning “construction,” “academic identity,” “authorial identity,” “plagiarism,” and “gender.” Figure 5 shows that most related burst keywords in this field appeared but lasted only briefly, except for “academic voice.” No burst keywords have emerged recently, indicating research in this field generally lacks vitality.

As Bruce (2008) indicated, effective academic writing required appropriate positioning within the academic community and the ability to display an identity as a scholar. It can be achieved through the construction of an authorial identity. On the other hand, individual academics constantly construct and present their identity as professional “selves” through academic writing. French (2020) pointed out that constructing and maintaining a positive academic (professional) identity were partly, yet significantly, achieved through professional writing in higher education habitus. Various social-cultural identities could influence academic writing (Belcher, 2009), and gender is the most discussed.



FIGURE 2
Chronological distribution of publication.

Keywords “academic identity,” “authorial identity,” and “plagiarism” are closely related to each other. A robust academic and well-established authorial identity is likely to foster a strong commitment to academic integrity and originality. Improving students’ authorial and academic identity contributes to reducing unintentional plagiarism as they can understand the role of the author better and take a more authorial role in their academic writing.

3.3 Visualization of clusters

The cosine index can represent the co-occurrence intensity: the larger the cosine value, the greater the co-occurrence intensity between keywords. This study applied the clustering principles proposed by Callon et al. (1991) to divide the clusters. With the aid of the *CiteSpace* software, a 272*272 keyword matrix was generated. The pair of keywords with the highest cosine index were identified as the theme words in the first cluster. The 272 keywords in the matrix were then sorted in descending order based on their cosine index with either of the keywords. Keywords with non-zero values were selected in descending order, including theme words, and the cluster name was summarized according to the content of the keywords within the cluster. If a cluster contains more than 10 keywords or less than 2 keywords, it will be excluded from being classified as a cluster. After generating a cluster, the keywords within this cluster were removed by deleting the corresponding rows and columns in the matrix, preventing these keywords from being included in subsequent clusters. Those steps were repeated until all keywords with co-occurrence relationships had been clustered (all remaining keywords had a co-occurrence intensity of 0). Forty-one clusters were identified, with 2 excluded because the number of members did not meet the required standards. Finally, 39 valid clusters were obtained through this process.

The average frequency of keywords within a cluster, minus the average frequency of all keywords, represents the attention level of the cluster. Similarly, the average initial year of appearance for

keywords within a cluster, minus the average initial year of appearance for all keywords, represents the novelty level of the cluster. A strategic coordinate diagram was plotted, with attention on the x-axis and novelty on the y-axis, as shown in Figure 6. The names and members of all clusters are listed in Supplementary material.

Clusters in the first quadrant have novelty and attention values greater than 0, suggesting that the research contents represented by these clusters are highly regarded and constitute current hotspots in this field.

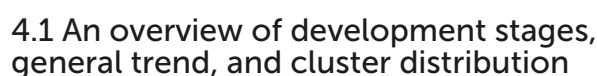
Clusters in the second quadrant have a novelty value greater than 0 but an attention value less than 0, indicating that related research contents are novel but have not yet gained the widespread attention they deserve. They are potential future hotspots and will gradually shift to the first quadrant, becoming more established research hotspots as interest increases.

Clusters in the third quadrant have both novelty and attention values less than 0, indicating that their research contents have drawn low attention due to their low novelty, placing them in marginalized research areas. These fields are now somewhat outdated.

Clusters in the fourth quadrant have an attention value greater than 0 but a novelty value less than 0, indicating that the contents represented by these clusters are well-regarded but not the recent research hotspots belonging to foundational research areas.

Through meticulous manual analysis of the collected articles, the clusters were further categorized into four fields: plagiarism related to authorial and academic identity; academic (professional) development related to authorial and academic identity; socio-cultural identity in academic writing; pedagogy of identity construction in academic writing.

In the following discussion section, a comprehensive and in-depth analysis of the clusters and their research contents will address each theme’s specific state of research, its characteristics, and practical implications, particularly regarding plagiarism prevention, pedagogy, and scholarly development. Greater attention will be given to the first and second quadrants because they contain clusters with a high level of novelty and represent current hotspots and substantial potential ones for further research.



Since 2017, the literature volume increased sharply, reaching its peak in 2020. With the development of higher education, the “publish or perish” culture has intensified (Nygaard, 2017), leading to unprecedented attention on academic writing and identity construction. As a result, a substantial body of literature has rapidly emerged and expanded in this area. This trend aligns with the large number of clusters in the third quadrant, which is the second largest group with nine clusters, although they now appear somewhat outdated. They are *academic voice, transitioning from professional work and Master’s coursework to the research dissertation, decolonization, socio-symbolic function of academic language, resistance by L2 writers, writing center from the consultant’s perspective rather than students,*

After 2020, it showed a sharp decline. This finding corresponds to the number of clusters in the first quadrant, which is the fewest, with only three, indicating that there are few current research hotspots in this field and that overall research activity is presently low. Previous research hotspots (in the third quadrant) have become outdated, while new emerging hotspots (in the second quadrant) remain in a latent phase and have yet to materialize fully. As a result, the volume of literature during this period sharply declined. However, it is not at the bottom, consistent with the large number of potential clusters in the second quadrant. Future research on identity may eventually return to a more stable developmental status as emerging hotspots transfer into current ones.

The majority of clusters are distributed in the second quadrant, with 20 clusters accounting for approximately 49% of the total. They currently represent potential future hotspots and related topics that need further exploration. The next two sections will elaborate on the current hotspots in the first quadrant and potential ones in the second quadrant, which require particular attention.

Three clusters are identified as current hotspots: one related to the field of plagiarism and the other two to sociocultural identity. First, plagiarism is an ethical violation involving using another's work or ideas without proper attribution, which undermines authorial and academic identity by violating academic integrity norms. Constructing a strong authorial and academic identity can, in turn, help prevent plagiarism. The current hotspot related to plagiarism is *academic*

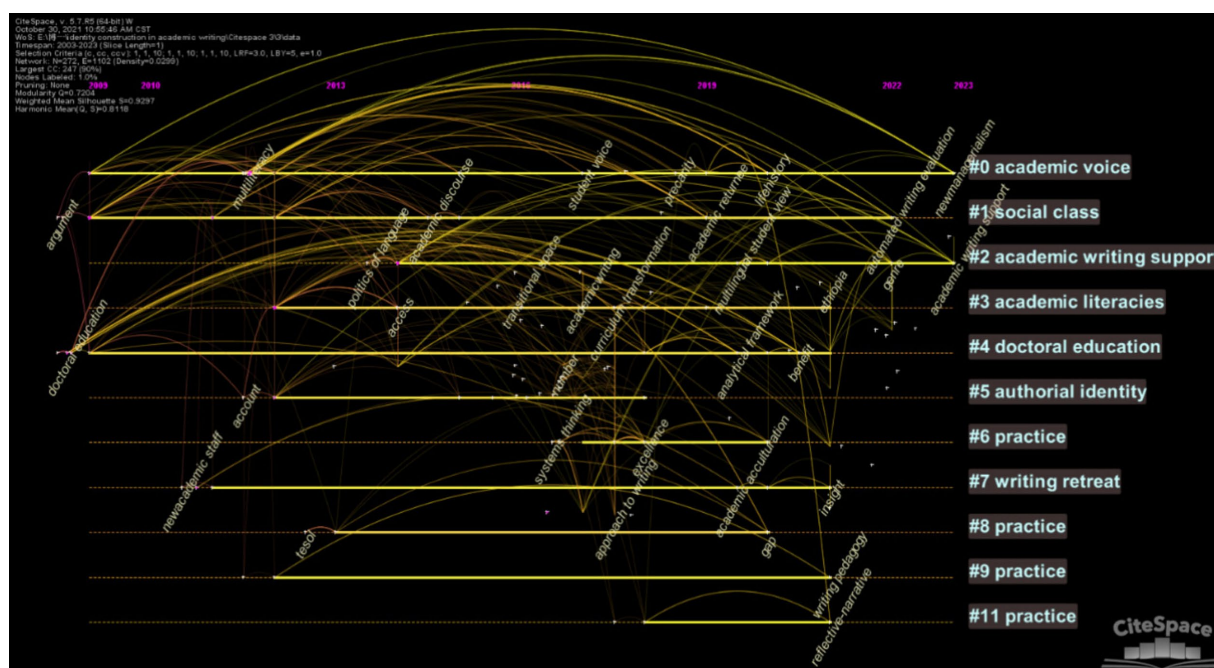


FIGURE 4
Keywords timeline visualization.

integrity, which has been studied from the perspective of Academic Integrity Socialization. Merkel (2022) found that most students had a limited and traditional understanding of plagiarism, viewing it solely as a violation. They were more focused on cultivating a moral academic identity as writers than on adhering to the academic community's formal guidelines and norms related to plagiarism. The perspective of Academic Integrity Socialization can address this problem to a certain degree. According to this perspective, students should be provided with a safe and supportive space to explore the *academic integrity* expectations of their institution in a comprehensive and learner-centered manner. Khoo and Kang (2022) investigated undergraduates' responses to Academic Integrity Socialization. They concluded that when students saw *academic integrity* as integral to their identity as junior scholars, they engaged more meaningfully in scholarly conversations in these spaces and contributed to the academic community with integrity and respect.

Second, the other two current hotspots are related to socio-cultural identity in academic writing. They are *academic literacy* and *English-as-a-second-language discourse*. *Academic literacy* is novel and popular with the highest burst strength, allowing scholars to examine issues of voice and writer identity in academic writing, which become more complex in multilingual contexts (Robinson-Pant and Wolf, 2017, p. 11). The geopolitics of academic publishing is one of the topics in this field. Through the lens of *academic literacy*, it was found that academics outside the Global North valued the opportunity to publish internationally to gain a voice in the global academic community. International journals are perceived to reach a larger audience, provide more rigorous peer review, and operate more efficiently than local journals (Getahun et al., 2021). *Academic literacy* also provides a theoretical foundation for creating socio-cultural spaces for writing support in alignment with the spatial nature of academic writing. Writing groups, increasingly popular in universities,

offer students a space to develop their voice and identity as academic writers (Aitchison and Lee, 2006; Larcombe et al., 2007). This conclusion was confirmed further by Wilmot and McKenna (2018). In addition to writing support forms like writing groups, researchers also identified other ways to enhance academic writing from the perspective of *academic literacy*, such as the appropriate use of various writing genres (Roald et al., 2021). In terms of the scope of application, the academic literacy model should be integrated into the practices of all students, regardless of their linguistic identities (Hathaway, 2015).

Another current hotspot is *English-as-a-second-language discourse*. It explores how identities and voices are formed in the increasingly prevalent transnational environments of today's globalized society, focusing on English academic discourse in higher education, particularly as a second language in plurilingual or multilingual writing contexts (Langum and Sullivan, 2020). This section presents the topics currently attracting the most attention from scholars, while the next section will elaborate on potential hotspots.

4.3 Potential hotspots

Twenty potential hotspots are distributed across the four fields of plagiarism, academic (professional) development, socio-cultural identity, and the pedagogy of identity construction. First, the potential future hotspots related to plagiarism focus on evaluating identity, including *measures of identity* and *discoursal identity*. Pittam et al.'s (2009) six-factor model (SAQ) and Ballantine et al.' (2013) three-factor model (alternative SAQ) are pioneering *measures* of evaluation. Based on them, scholars continued to refine the measures, illustrate the discursive embodiment of the measures, and elaborate attributes related to authorial identity (Cheung et al., 2017; Cheung et al., 2018;

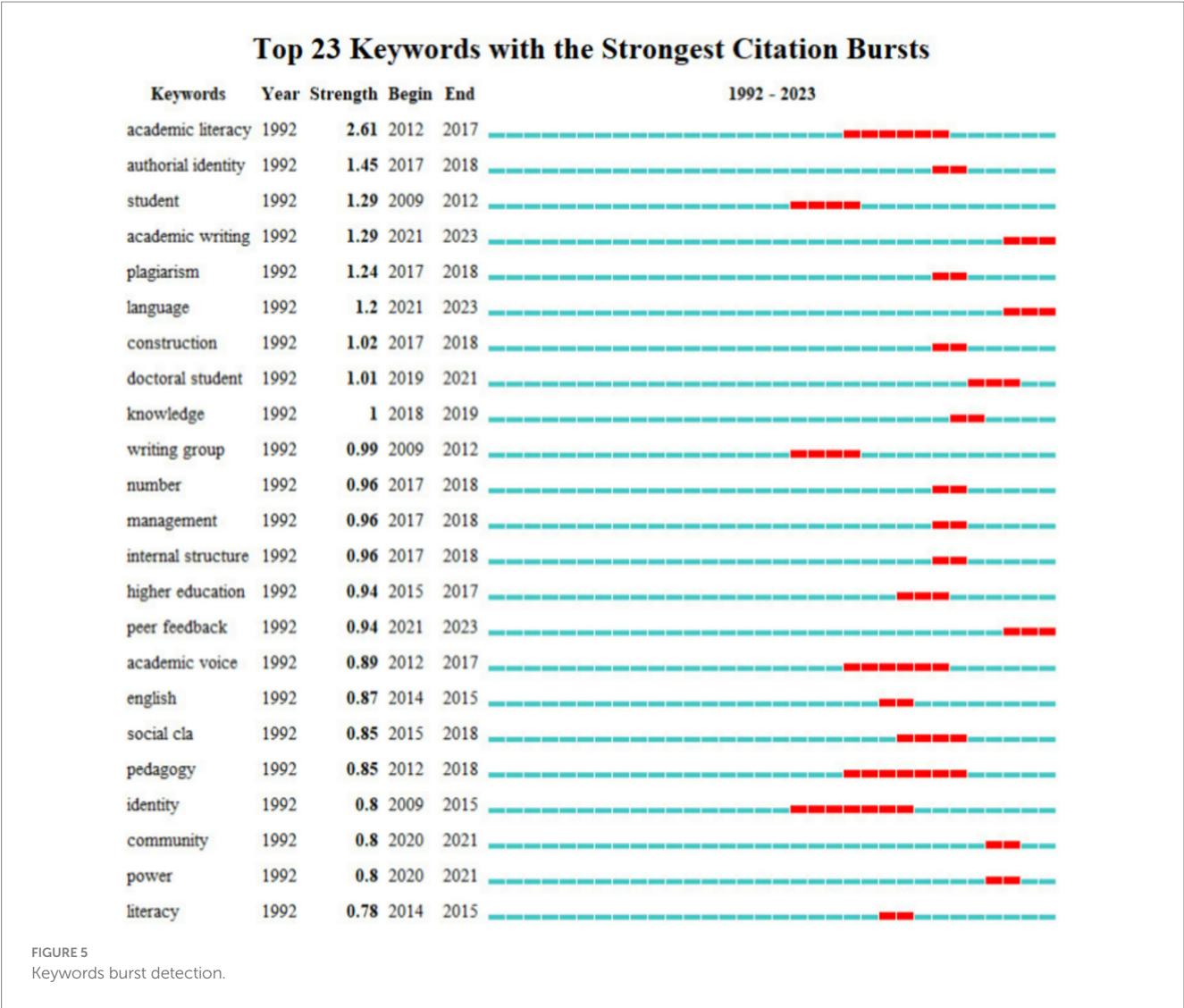


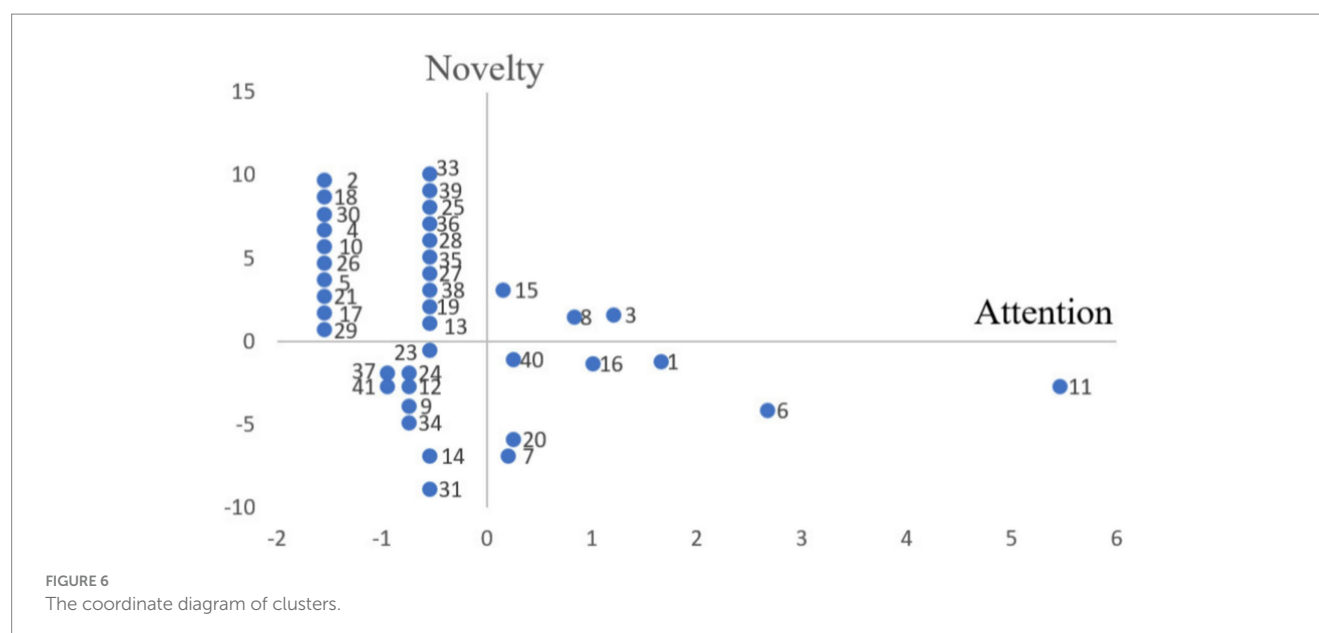
TABLE 1 Keywords with a frequency of 3 or more.

Keyword	Frequency	The initial year	Keyword	Frequency	The initial year
Academic literacy	12	2012	Plagiarism	4	2017
Student	8	2009	Academic identity	3	2011
Literacy	7	2014	Teacher	3	2011
English	6	2014	Doctoral student	3	2019
Construction	5	2009	Authorial identity	3	2017
Discourse	5	2012	Language	3	2021
Pedagogy	5	2012	Multiliteracy	3	2012
University	5	2017	Doctoral education	3	2009
Writing group	4	2009	Gender	3	2009

Kanwal et al., 2021). Some scholars have examined these measures in the context of non-English speakers, particularly Chinese students (Ballantine et al., 2018), because English academic writing is becoming increasingly important for multilingual students with the globalization of higher education and the academic community. Preventing plagiarism by enhancing authorial and academic identities is a current

hotspot, but its effectiveness depends on solid and reliable evaluation measures (Cheung et al., 2017). Therefore, these measures appear as potential hotspots.

Second, five potential hotspots were identified within the academic (professional) development field. On one hand, they focus on the spatial nature of academic writing practice (Beighton, 2020),



which aligns with the current hotspot *academic literacy*. Both confirm the importance of a safe and supportive space for promoting academic writing and identity, as students often face alienation and fear accusations of plagiarism due to limited English proficiency, which can negatively impact their relationships with peers and instructors, as well as their sense of belonging in the academic community. Creating such a space is crucial for addressing this issue. This *spatial nature* can be embodied in various forms, such as writing groups and retreats. Writing groups, as spaces for academic writing development, offered a transformative framework that supported proactive student learning through peer interaction (Wilmot and McKenna, 2018). Writing retreats could alleviate the isolation associated with academic writing, thereby improving scholars' sense of belonging within the academic community and leading to better hedonic and eudaimonic *well-being* for writers when the interventions were sustained (Eardley et al., 2021).

On the other hand, they focus on *the spatial dynamics and transition of academic identity*, particularly among returnees to China from abroad, doctoral students, and educators, because these are the three most representative groups to undergo a spatial transition in the academic community of higher education. The journey to becoming a mature academic is ongoing, as it requires adapting to evolving writing standards and the changing policies of institutions across different regions as needed. The norms and values of academic writing are acquired, negotiated, and sometimes resisted during the transition of doctoral identity to academics (Katila et al., 2019). Doctoral graduates who study abroad and return to work in Chinese universities face significant challenges in *(re)constructing academic identity as returnees*, particularly when adapting to various academic assessment policies in Chinese higher education (Ai, 2019). For educators, academic identity can be developed through the repeated transitions between researcher and teacher roles when engaging in systematic study or teaching and learning practices. The transition is also beneficial for writing *SoTL (scholarship of teaching and learning) identity* articles for publication, which can lead to greater recognition within the academic community (Healey et al., 2019).

Third, the nine potential hotspots related to sociocultural identity can further be divided into two themes: linguistic factors resulting from the globalization of the academic community and other social factors such as individual experiences, nationality, gender, and ethnicity. The influence of *bilingualism/multilingualism* in constructing multiple identities as an academic writer has been confirmed (Asadolahi and Nushi, 2021). *Translanguaging* originated in bilingual education and was an adaptable strategy to utilize the entire linguistic repertoire during the learning process. It has been proven effective for promoting identity construction when introduced in English for Academic Purposes (EAP) courses (Hiller, 2021). However, the strategy of *translanguaging* should be employed judiciously, as multilingual students perceive *language mixing* in formal writing as inappropriate and potentially undermining their academic identity (Kafle, 2020). Therefore, they tend to restrict their *multilingual* and *multiliterate* skills to conform to institutional standards for conventional academic writing in formal English academic contexts while negotiating their multiple identities in practice (Marshall et al., 2012). In response to this finding, Gagne et al. (2023) suggested that educators could create more positive and inclusive academic environments by recognizing and utilizing students' diverse linguistic abilities rather than viewing them as challenges to be overcome, to make them effectively use their *multilingual* and *multiliterate* skills. In addition to academic environments, multilingual proficiency and sociocultural identities can influence students' academic writing engagement. Students with a more positive sociocultural identity and mastery-oriented learning beliefs are more actively involved and can progress more (Zhang and Xu, 2022).

Some other individual and social factors can potentially become hotspots in the future. *Life experience* is one of them because the depth of personal experiences shapes various roles involved in academic identity, such as creator, interpreter, communicator, and presenter (Lo et al., 2020). As a result, personal experience also influences academic writing by shaping the writers' *self-perception* (Clark and Ivanic, 2013). It supports the finding that writing projects are most meaningful when students have opportunities to connect their writing to personal factors, including the sense of authorship and vision for future writing

or identities (Eodice et al., 2019). This explains why *shifting the focus from the final product to the writing process* is a prevalent trend in academic writing pedagogy, as it better leverages students' experiences by encouraging and sustaining students' agency and identity in higher education. In addition, gaining a deeper understanding of *gender*, *social class*, *nationality*, and *ethnicity* identities further helps students navigate their approach to academic writing more effectively (Preece, 2018; Danvers et al., 2019; Zhang and Xu, 2022; Pham and Hayden, 2019).

Fourth, there are four potential future hotspots in identity construction pedagogy. Three focus on *doctoral students* and education, further corroborating the findings from the keyword co-occurrence analysis. The function of *mindfulness practices* in doctoral academic writing has been confirmed because they can help participants better understand the writing process and writer identity, accepting both themselves and others as writers through self-reflection, creativity, and joy in writing (Woloshyn et al., 2022). This conclusion is supported by the process of *writing a doctoral thesis*, which includes key elements such as selecting a topic/title for the project, writing the abstract, conducting the literature review, and performing the analysis. Throughout this process, interactions with support networks play a crucial role in shaping, developing, and refining academic identity (Nartey, 2021). Besides, it is proposed that *doctoral student performance* should be evaluated based on the dimension of academic writing (Ward and Brennan, 2020). Specifically, Ward and Brennan (2020) added subdimensions of student-learning identity fit and student-(academic) writing fit to extend Baker and Pifer's (2015) multidimensional framework of student-doctoral fit. The new model also serves as a tool to develop instruments, such as surveys or assessments, to test specific hypotheses or propositions about *doctoral student performance*.

This study identifies three dimensions of identity involved in academic writing, facilitating a more comprehensive theoretical understanding of how identity is reflected, negotiated, and constructed in academic writing contexts. By employing the strategic coordinate diagram, it visualizes both the current research hotspots and underexplored areas in this field, enabling researchers to understand the current state of research and identify gaps and opportunities for further inquiry, therefore advancing the theoretical landscape.

5 Conclusion

5.1 Major findings

The literature in this field typically showed a growing trend before 2020 but declined steeply after 2020. The years 2016 and 2020 marked significant milestones in its development. Published articles exhibited a modest increase despite some fluctuations before 2016. Since 2017, it has risen sharply and reached its peak in 2020. After that, it saw a steep decline up until 2022.

Only three clusters can be regarded as current research hotspots. They are distributed in plagiarism and socio-cultural identity and focus on academic integrity, academic literacy, and English-as-a-second-language discourse. The majority of clusters, 20 in total, are identified as potential future hotspots. They are distributed across all four fields: plagiarism, academic (professional) development, socio-cultural identity, and the pedagogy of identity construction.

Potential future hotspots in plagiarism focus on measures of authorial and academic identity, while those in the academic (professional) development field highlight the spatial aspects of academic writing and explore ways to promote academic growth and well-being for various groups in higher education, including doctoral students, instructors, and returnees. The potential future hotspots in the field of socio-cultural identity center on how linguistic and social identities influence academic writing in the context of the globalized academic community. Linguistic factors involve *translanguaging*, *language mixing*, *multiliteracy*, *English language privilege*, and *bilingualism/multilingualism*. Other social-cultural factors include *individual experiences*, *nationality*, *gender*, and *ethnicity*.

Potential future hotspots in the pedagogy of identity construction highlight practical approaches to enhancing identity development and academic writing. Universities are suggested to support *mindfulness practices*, recognize academic identity in writing a *doctoral thesis*, highlight the *process of writing* rather than the final product, and consider academic writing, grounded in a strong sense of identity, as a key *assessment of doctoral student performance*.

Most of the research in this field has been conducted using thematic analysis as a qualitative method. Most clusters are distributed in the second quadrant, indicating a generally high level of novelty, though they have not received the attention they deserve. The second-largest group of clusters is located in the third quadrant, reflecting somewhat outdated topics. Therefore, most studies in this field either exhibit high novelty but lack attention as future hotspots or are overlooked due to being outdated. The current hotspots are relatively rare.

Studies focusing on socio-cultural identity and pedagogy of identity construction are more thriving than the other two, with more clusters representing. However, clusters in the second quadrant related to academic (professional) development account for the most significant proportion, indicating their great potential to generally become hotspots in the future.

5.2 Implications

Enhancing authorial and academic identity is an effective way to prevent plagiarism. Students often view plagiarism solely as an ethical violation and fail to understand it within the broader framework of academic writing norms. Providing a safe and supportive space is beneficial for them to examine expectations for academic integrity from a comprehensive, detailed, and learner-centered perspective. Studies on effective and valid measures for evaluating authorial and academic identity offer valuable insights into its development in academic writing.

Based on the spatial nature of academic writing practice, writing supports, such as writing groups, centers, retreats, and mindfulness practice, effectively alleviate isolation and enhance scholars' sense of belonging within their academic identity, thereby promoting their development as academics.

In the context of socio-cultural identity in academic writing, writing groups can be organized from the perspective of academic literacy, with the academic literacy model integrated into the practices of all students, regardless of their linguistic backgrounds.

Educators should recognize and value students' multilingual abilities to develop inclusive teaching methods that better support

non-native students in English-dominant academic environments. Additionally, the construction of academic identity should be incorporated into doctoral theses, with the focus shifting from the final academic product to the writing process. Authorial identity should be treated as a form of tacit knowledge to be developed, evolving through maturing and gaining experience as a writer.

5.3 Limitations

The study is far from flawless, and several limitations must be acknowledged. First, the *CiteSpace analysis* was confined to data derived from the Web of Science (WoS) database due to the data format requirements. While the results are basically representative, they do not capture the full breadth of existing literature. Future research should integrate data from additional databases such as Scopus and ScienceDirect for a more comprehensive analysis.

Second, the names and content of the clusters may have overlapping meanings with less distinct boundaries, as a single publication can address multiple specific themes. Consequently, cluster names primarily reflect the main themes expressed, potentially overlooking the nuanced interconnections between different topics.

Third, the exclusion of studies published in non-English languages inevitably introduces a bias into the analysis of this important topic, thereby reinforcing the dominance of English in research on multilingual writers' practices. Addressing these limitations in future studies will strengthen the robustness and applicability of the research outcomes.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

References

- Ai, B. (2019). Pains and gains of working in Chinese universities: an academic returnee's journey. *High. Educ. Res. Dev.* 38, 661–673. doi: 10.1080/07294360.2019.1590320
- Aitchison, C., and Lee, A. (2006). Research writing: problems and pedagogies. *Teach. High. Educ.* 11, 265–278. <https://doi.org/10.1080/13562510600680574>. doi: 10.1080/13562510600680574
- Archer, A., and Parker, S. (2016). Transitional and transformational spaces: mentoring young academics through writing centers. *Educ. Change* 20, 43–58. doi: 10.17159/1947-9417/2016/567
- Asadolahi, S., and Nushi, M. (2021). An autobiographical narrative inquiry of academic identity construction of PhD candidates through L2 authorship development. *Heliyon* 7:e06293. doi: 10.1016/j.heliyon.2021.e06293
- Baker, V. L., and Pifer, M. J. (2015). Antecedents and outcomes: theories of fit and the study of doctoral education. *Stud. High. Educ.* 40, 296–310. doi: 10.1080/03075079.2013.823936
- Ballantine, J., Guo, X., and Larres, P. (2013). Psychometric evaluation of the student authorship questionnaire: a confirmatory factor analysis approach. *Stud. High. Educ.* 40, 596–609. doi: 10.1080/03075079.2013.835910
- Ballantine, J., Guo, X., Larres, P., and Yu, M. (2018). Chinese authorial identity: a model for scoring the student authorship questionnaire. *Stud. High. Educ.* 43, 2385–2397. doi: 10.1080/03075079.2017.1333493
- Beighton, C. (2020). Beyond alienation: spatial implications of teaching and learning academic writing. *Teach. High. Educ.* 25, 205–222. doi: 10.1080/13562517.2018.1554642
- Belcher, D. D. (2009). How research space is created in a diverse research world. *J. Second. Lang. Writ.* 18, 221–234. doi: 10.1016/j.jslw.2009.08.001
- Botelho De Magalhães, M., Cotterall, S., and Mideros, D. (2019). Identity, voice and agency in two EAL doctoral writing contexts. *J. Second. Lang. Writ.* 43, 4–14. doi: 10.1016/j.jslw.2018.05.001
- Bruce, I. (2008). "Theorising tertiary writing instruction: accounting for the process, postprocess, genre and critical literacies approaches" in Proceedings of the TWN biennial colloquium: From Here to there (Auckland: Auckland University of Technology), A1–A8.
- Callon, M., Courtial, J. P., and Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemistry. *Scientometrics* 22, 155–205. doi: 10.1007/BF02019280
- Canagarajah, A. S. (2015). "Blessed in my own way:" pedagogical affordances for dialogical voice construction in multilingual student writing. *J. Second. Lang. Writ.* 27, 122–139. doi: 10.1016/j.jslw.2014.09.001
- Canagarajah, A. S. (2020). Transnational literacy autobiographies as translingual writing. New York: Routledge.
- Chen, C. M. (2006). CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inf. Sci. Technol.* 57, 359–377. doi: 10.1002/asi.20317
- Cheung, K. Y. F., Elander, J., Stupple, E. J. N., and Flay, M. (2018). Academics' understandings of the authorial academic writer: a qualitative analysis of authorial identity. *Stud. High. Educ.* 43, 1468–1483. doi: 10.1080/03075079.2016.1264382
- Cheung, K. Y. F., Stupple, E. J. N., and Elander, J. (2017). Development and validation of the student attitudes and beliefs about authorship scale: a psychometrically robust measure of authorial identity. *Stud. High. Educ.* 42, 97–114. doi: 10.1080/03075079.2015.1034673

Author contributions

YT: Conceptualization, Investigation, Methodology, Visualization, Writing – original draft. DL: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1499917/full#supplementary-material>

- Clark, R., and Ivanic, R. (2013). *The politics of writing*. New York: Routledge.
- Danvers, E., Hinton-Smith, T., and Webb, R. (2019). Power, pedagogy and the personal: feminist ethics in facilitating a doctoral writing group. *Teach. High. Educ.* 24, 32–46. doi: 10.1080/13562517.2018.1456423
- Eardley, A. F., Banister, E., and Fletcher, M. (2021). Can academic writing retreats function as wellbeing interventions? *J. Furth. High. Educ.* 45, 183–196. doi: 10.1080/0309877X.2020.1744542
- Eodice, M., Ellen Geller, A., and Lerner, N. (2019). The power of personal connection for undergraduate student writers. *Res. Teach. Engl.* 53, 320–339. doi: 10.58680/rte201930141
- French, A. (2020). Academic writing as identity-work in higher education: forming a 'professional writing in higher education habitus'. *Stud. High. Educ.* 45, 1605–1617. doi: 10.1080/03075079.2019.1572735
- Gagne, A., McIntosh, M., Herath, S., Fowler, M.-A., Kim, J., Baxan, V., et al. (2023). Let's talk about writing support for plurilingual graduate students: a collaborative autoethnography. *TESL Canada J.* 40, 81–106. doi: 10.18806/tesl.v40i1/1386
- Getahun, D. A., Hammad, W., and Robinson-Pant, A. (2021). Academic writing for publication: putting the 'international' into context. *Res. Comp. Int. Educ.* 16, 160–180. doi: 10.1177/17454999211009346
- Gourlay, L. (2011). New lecturers and the myth of 'communities of practice'. *Stud. Contin. Educ.* 33, 67–77. doi: 10.1080/0158037X.2010.515570
- Hathaway, J. (2015). Developing that voice: locating academic writing tuition in the mainstream of higher education. *Teach. High. Educ.* 20, 506–517. doi: 10.1080/13562517.2015.1026891
- Healey, M., Matthews, K. E., and Cook-Sather, A. (2019). Writing scholarship of teaching and learning articles for peer-reviewed journals. *Teach. Learn. Inq.* 7, 28–50. doi: 10.20343/teachlearningqu.7.2.3
- Hiller, K. E. (2021). Introducing translanguaging in an EAP course at a joint-venture university in China. *RELJ* 52, 307–317. doi: 10.1177/00336882211014997
- Ivanić, R. (1998). *Writing and identity: The Discoursal construction of identity in academic writing*. Amsterdam: John Benjamins Publishing Company.
- James, B. (2012). Becoming an 'authorised' postgraduate research writer. *Innov. Educ. Teach. Int.* 49, 41–50. doi: 10.1080/14703297.2012.647782
- Kafle, M. (2020). "No one would like to take a risk": multilingual students' views on language mixing in academic writing. *System* 94:102326. doi: 10.1016/j.system.2020.102326
- Kanwal, N., Amin Qadir, S., and Shaikat, K. (2021). Writing instructions at a university and identity issues: a systemic functional linguistics perspective. *Int. J. Emerg. Technol. Learn.* 16:275. doi: 10.3991/ijet.v16i06.19485
- Katila, S., Laamanen, M., Laihonon, M., Lund, R., Meriläinen, S., Rinkinen, J., et al. (2019). Becoming academics: embracing and resisting changing writing practice. *Qualitat. Res. Organ. Manag. Int. J.* 15, 315–330. doi: 10.1108/QROM-12-2018-1713
- Khoo, E., and Kang, S. (2022). Proactive learner empowerment: towards a transformative academic integrity approach for English language learners. *Int. J. Educ. Integr.* 18:24. doi: 10.1007/s40979-022-00111-2
- Langum, V., and Sullivan, K. P. H. (2020). Academic writing, scholarly identity, voice and the benefits and challenges of multilingualism: reflections from Norwegian doctoral researchers in teacher education. *Linguist. Educ.* 60:100883. doi: 10.1016/j.linged.2020.100883
- Larcombe, W., McCosker, A., and O'Loughlin, K. (2007). Supporting education PhD and DEd students to become confident academic writers: an evaluation of thesis writers' circles. *J. Univ. Teach. Learn. Pract.* 4, 61–71. doi: 10.53761/1.4.1.6
- Lea, M. R., and Street, B. V. (2006). The "academic literacies" model: theory and applications. *Theory Pract.* 45, 368–377. doi: 10.1207/s15430421tip4504_11
- Liu, P.-H. E., and Tannacito, D. J. (2013). Resistance by L2 writers: the role of racial and language ideology in imagined community and identity investment. *J. Second. Lang. Writ.* 22, 355–373. doi: 10.1016/j.jslw.2013.05.001
- Lo, Y. Y., Othman, J., and Lim, J. W. (2020). Multiplex aspects in the construction of academic writer identity among ESL doctoral students. *3L The Southeast Asian J. Engl. Lang. Stud.* 26, 110–123. doi: 10.17576/3L-2020-2603-09
- Marshall, S., Hayashi, H., and Yeung, P. (2012). Negotiating the multi in multilingualism and multiliteracies: undergraduate students in Vancouver, Canada. *Can. Mod. Lang. Rev.* 68, 28–53. doi: 10.3138/cmlr.68.1.028
- Merkel, W. (2022). Simple, yet complex: pre-service teachers' conceptions of plagiarism at a Norwegian university. *Scand. J. Educ. Res.* 66, 923–935. doi: 10.1080/00313831.2021.1939778
- Mitchell, K. M. (2017). Academic voice: on feminism, presence, and objectivity in writing. *Nurs. Inq.* 24:e12200. doi: 10.1111/nin.12200
- Nartey, M. (2021). Navigating the doctorate: a reflection on the journey of 'becoming' a PhD in applied language sciences. *Cogent Educ.* 8:1994360. doi: 10.1080/2331186X.2021.1994360
- Nygaard, L. P. (2017). Publishing and perishing: an academic literacies framework for investigating research productivity. *Stud. High. Educ.* 42, 519–532. doi: 10.1080/03075079.2015.1058351
- Okuda, T., and Anderson, T. (2018). Second language graduate students' experiences at the writing center: a language socialization perspective. *TESOL Q.* 52, 391–413. doi: 10.1002/tesq.406
- Pham, L. T., and Hayden, M. (2019). Research in Vietnam: the experience of the humanities and social sciences. *J. Int. Compar. Educ.* 8, 27–40. doi: 10.14425/jice.2019.8.1.27
- Pittam, G., Elander, J., Lusher, J., Fox, P., and Payne, N. (2009). Student beliefs and attitudes about authorial identity in academic writing. *Stud. High. Educ.* 34, 153. https://doi:10.1080/03075070802528270–170. doi: 10.1080/03075070802528270
- Preece, S. (2018). Identity work in the academic writing classroom: where gender meets social class. *J. Engl. Acad. Purp.* 32, 9–20. doi: 10.1016/j.jeap.2018.03.004
- Raitskaya, L., and Tikhonova, E. (2022). An in-depth glimpse into research on academic writing. *J. Lang. Educ.* 8, 5–16. doi: 10.17323/jle.2022.14586
- Roald, G. M., Wallin, P., Hybertsen, I. D., and Stenoien, J. M. (2021). Learning from contrasts: first year students writing themselves into academic literacy. *J. Furth. High. Educ.* 45, 758–770. doi: 10.1080/0309877X.2020.1813264
- Robinson-Pant, A., and Wolf, A. (2017). *Researching across languages and cultures: A guide to doing research Interculturally*. London: Routledge.
- Shaw, C., and Le Roux, K. (2017). From practitioner to researcher: designing the dissertation process for part time coursework masters students. *Syst. Pract. Action Res.* 30, 433–446. doi: 10.1007/s11213-016-9402-7
- Ward, A. M., and Brennan, N. M. (2020). Developing a student-doctoral education fit analytical model to assess performance. *Stud. High. Educ.* 45, 1448–1460. doi: 10.1080/03075079.2018.1545758
- Wilmot, K., and McKenna, S. (2018). Writing groups as transformative spaces. *High. Educ. Res. Dev.* 37, 868–882. doi: 10.1080/07294360.2018.1450361
- Woloshyn, V., Obradović-Ratković, S., Julien, K., Rebek, J.-L., and Sen, A. P. (2022). Breathing our way into mindful academic writing: a collaborative autoethnography of an online writing community. *J. Furth. High. Educ.* 46, 1135–1148. doi: 10.1080/0309877X.2022.2055450
- Zhang, Z., and Xu, L. (2022). Student engagement with automated feedback on academic writing: a study on Uyghur ethnic minority students in China. *J. Multiling. Multicult. Dev.* 45, 3466–3479. doi: 10.1080/01434632.2022.2102175



OPEN ACCESS

EDITED BY

Xinghua Liu,
Shanghai Jiao Tong University, China

REVIEWED BY

David Pérez-Jorge,
University of La Laguna, Spain
Zhi Yang,
Griffith University, Australia
Mark Feng Teng,
Macao Polytechnic University, Macao SAR,
China
Liyong Nong,
Hezhou University, China
Donghong Liu,
Southeast University, China

*CORRESPONDENCE

Ruoxi Liu
✉ ruoxiliu06060@163.com
Ping Xin
✉ xinpingskuedu@163.com

RECEIVED 12 July 2024
ACCEPTED 09 December 2024
PUBLISHED 07 January 2025

CITATION

Liu R and Xin P (2025) Self-imposed pressure or joyful learning: emotions of Chinese as a foreign language learners in feedback on academic writing.
Front. Psychol. 15:1463488.
doi: 10.3389/fpsyg.2024.1463488

COPYRIGHT

© 2025 Liu and Xin. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Self-imposed pressure or joyful learning: emotions of Chinese as a foreign language learners in feedback on academic writing

Ruoxi Liu* and Ping Xin*

School of Chinese as a Second Language, Faculty of Humanities, Peking University, Beijing, China

Although writing feedback is widely believed to elicit a range of emotions, studies on the emotional experiences of L2 students with this teaching and learning tool, as well as their regulation strategies, remain largely underexplored. Drawing on the analytical framework of academic emotions from the perspective of positive psychology, this study examines two Chinese as foreign language (CFL) students' emotional reactions to their teacher's oral and written feedback and their emotion regulation strategies. The main data includes interviews, retrospective oral reports, students' reflection journals, academic writings, and teacher feedback. The study found that feedback aroused students' academic achievement emotions, cognitive emotions, and social emotions across various dimensions of valence and activation. Over the course of three feedback processes within one semester, the two learners' emotions gradually became neutral or positive. They effectively employed emotion-oriented, appraisal-oriented, and situation-oriented strategies to manage negative emotions and adapt to feedback. The findings suggest that paying attention to the intrinsic values of feedback may help learners experience more positive academic emotions, while paying too much attention to its extrinsic values may lead to negative emotions.

KEYWORDS

positive psychology, CFL academic writing, feedback, academic emotion, emotion regulation strategies

Introduction

The process of learning academic writing is filled with difficulties and challenges, especially for novice second-language writers who inevitably experience a range of complex emotions (Han and Hyland, 2019). Feedback, as a critical tool in writing instruction, refers to written or oral comments and revision suggestions on the language or content of a learner's writing. It plays a pivotal role in shaping the quality of second-language learners' writing output (Kepner, 1991). Learners are the primary agents of feedback information processing (Winstone et al., 2022), and the effectiveness of this processing directly impacts their writing abilities.

As an essential part of feedback and academic writing learning (Hyland and Hyland, 2006), emotional experiences significantly affect learners' L2 learning achievements, feedback engagement, and the effectiveness of their processing (Burić et al., 2016). Many teachers advocate for providing positive feedback while selectively providing negative feedback to reduce learners' negative emotions (Han and Hyland, 2019), enhance their learning behavior and sense of agency, and improve their academic achievement.

In recent years, positive psychology has led to the emotional turn in second language acquisition, and more and more scholars have called for attention to both individual learning outcomes and their emotional experiences (Dewaele and Li, 2020). However, research on the

emotional experiences of second language learners in response to writing feedback is still insufficient; most existing studies focus on single writing feedback (Han and Xu, 2020; Geng and Yu, 2024), so it is impossible to observe the dynamic development of an individual's emotional experiences over a period of time after several feedback sessions and whether emotions can be effectively improved through regulatory strategies. In addition, previous studies have mostly assumed that learners with lower language proficiency and writing ability are more likely to receive more feedback, resulting in more negative emotions with higher activation levels being aroused (Hyland, 1998; Jiang and Dewaele, 2019). The specific relationship between learners' academic achievement and emotional experiences needs to be analyzed in more detail. Based on this, to advance our understanding of the emotional dimension of feedback, this study investigated the dynamic changes in the emotional experiences and regulation strategies of two undergraduate CFL students during a one-semester Chinese academic writing course evoked by three feedback sessions.

Literature review

Emotions typically include both trait emotions (habitual and recurring) and state emotions ("momentary occurrences within a given situation at a specific point in time") (Pekrun, 2006, p. 317). This study uses a qualitative approach to explore the emotional experiences of individual CFL students in response to feedback from a cognitive perspective. Therefore, we focus primarily on state emotions, which are understood as subjective academic emotions that are triggered by the moment when CFL students receive and process feedback.

Academic emotions and emotion regulation

Academic emotions refer to a series of emotions experienced by learners during academic activities. They are subjective psychological and physiological states that are directly related to academic learning, classroom teaching, and academic achievement (Pekrun et al., 2002; Imai, 2010). Among several theoretical approaches to the study of emotions (Gross, 1998; Han and Gao, 2024), Pekrun and his colleagues have provided a useful analytical framework for the study of academic emotions (Pekrun, 2006; Pekrun and Linnenbrink-Garcia, 2012). This framework considers multiple dimensions of emotions and is an analytical framework that has been widely used and recognized in the current field of academic writing to study L2 learners' emotions (e.g., Han and Hyland, 2019; Han and Xu, 2020; Geng and Yu, 2024). Pekrun (2006) research on academic emotions has conceptualized emotions as varying along two dimensions, i.e., valence and activation, and having different object foci. Valence refers to the positive-negative dimension of emotion, whereas activation refers to the degree of emotional arousal, which can be divided into activating and deactivating (Pekrun et al., 2002). On this basis, academic emotions are divided into positive activating emotions, positive deactivating emotions, negative activating emotions, and negative deactivating emotions. Although this framework has been widely applied to explore learners' emotional experiences, its establishment mainly relies on questionnaires. A single research method may not be able to

comprehensively and in-depth reveal individuals' emotional changes, nor can it disclose their differences based on individual experiences, personalities, etc. Moreover, the framework was mostly developed based on the Western cultural background, while learners from different cultural backgrounds may have different understandings and experiences of emotions. Therefore, the analytical frameworks of Pekrun (2006) and Pekrun and Linnenbrink-Garcia (2012) still need to be further verified and explored in different situations, cultural backgrounds, and individual emotional experiences of learners.

It is generally believed that positive emotions can expand learners' thought-action repertoire in the short term, which is conducive to building long-term cognitive, motivational, and social resources (Fredrickson, 2001), whereas excessive negative emotions may hinder their learning investment (Li et al., 2023; Han et al., 2024). However, academic emotions have the same valence but different activations that have different effects on academic achievement (Pekrun et al., 2002). For example, negative activating emotions (e.g., anxiety) may produce task-irrelevant thinking and undermine the students' intrinsic motivation, whereas negative deactivating emotions (e.g., hopelessness) could impair students' performance by undermining both intrinsic and extrinsic motivation, learning engagement, and promoting superficial information processing (Pekrun, 2006; Geng and Yu, 2024). In addition to valence and activation, Pekrun and Linnenbrink-Garcia (2012, p. 262) and Han and Hyland (2019) also grouped academic emotions according to their object focus. Table 1 categorizes academic emotions by object focus in feedback situations.

Regarding academic emotional experiences, Pekrun (2006) also proposed the control-value theory, which argues that emotions are mainly influenced by two factors: control appraisal and value appraisal. Control appraisal refers to learners' perceptions of their ability to control past, present, and future academic activities or outcomes. Value appraisal involves learners' judgments about the significance of academic activities or outcomes, which can be divided into intrinsic values and extrinsic values.

Intrinsic values refer to the subjective importance learners attach to an activity itself, regardless of its outcomes. For example, students who enjoy memorizing Chinese vocabulary and find it meaningful may value the activity even if it does not directly improve their Chinese learning performance. In contrast, extrinsic values view the activity as a means to achieve a specific outcome. For example, students might focus on improving their Chinese writing skills because it directly contributes to better scores on preparatory exams and increases their chances of entering undergraduate studies (Pekrun, 2006). Despite the importance of these factors, studies examining the influence of control and value appraisals on the emotional experiences induced by feedback remain scarce in current research (Han and Hyland, 2019).

Many studies have confirmed that individuals can regulate and cognitively process emotional experiences (Mahfoodh, 2017; Pitt and Norton, 2017; Geng and Yu, 2024). Academic emotion regulation is the process by which individuals influence the emotions they arouse when emotions occur and the expression of emotional experiences in a learning environment (Gross, 1998). Inspired by control value theory, Pekrun (2006) further proposed an analytical framework for academic emotion regulation strategies that include four emotion regulation strategies: emotion-oriented regulation, appraisal-oriented regulation, problem-oriented regulation, and situation-oriented regulation. Definitions of the emotion regulation strategies are given

TABLE 1 Academic emotions are categorized by object focus in feedback situations (Han and Hyland, 2019; Pekrun and Linnenbrink-Garcia, 2012).

Category	Definition	Definition in feedback situations
Achievement emotions		
-Activity emotions	Emotions are aroused when engaging in learning activities.	Emotions are aroused by processing and using feedback to improve the writing.
-Prospective outcome emotions	Emotions aroused by future expected outcomes.	Emotions are aroused by (a) the expected accuracy of the writing before receiving feedback and (b) the expected accuracy of the revised writing during revision.
-Retrospective outcome emotions	Emotions aroused by past task outcomes.	Emotions are aroused by (a) the written accuracy of the previous writing(s) after receiving feedback and (b) the written accuracy of the revised writing after revision.
Epistemic emotions	Emotions are aroused by cognitive processing during the task.	Emotions are aroused by the cognitive processing of feedback.
Social emotions	Emotions aroused by other persons.	Emotions are aroused by other persons, e.g., teachers and classmates, in feedback situations.
Topic emotions	Emotions aroused by the contents of learning material (e.g., empathy with the characters in a novel).	Emotions pertaining to the topic and content of the writing task.

TABLE 2 Academic emotion regulation strategies (Pekrun, 2006).

Strategies	Definitions
Emotion-oriented regulation	Regulate academic emotions directly (e.g., focusing attention on the emotion or distracting it away, using relaxation techniques, or taking drugs).
Appraisal-oriented regulation	Addresses the control and value antecedents of emotions (e.g., restructuring expectancies and attributions).
Problem-oriented regulation	Focuses on improving academic learning and achievement underlying perceived control (e.g., acquiring study skills).
Situation-oriented regulation	Attempt to change situational circumstances defining controllability and values (e.g., by asking for a reduction of task demands or by dropping out of a course).

in Table 2. This framework is currently widely used to study emotion regulation in the field of L2 academic writing (Han and Xu, 2020). Based on this framework, this study will further explore CFL students' emotion regulation strategies in feedback and explore whether these strategies can effectively improve their negative emotions.

Emotional experiences and emotion regulation in feedback on L2 academic writing

In the field of academic writing, previous studies have found that novice writers are more likely to experience multiple emotions, such as anxiety and pleasure, when receiving feedback (Lei and Hu, 2019; Yu and Jiang, 2022; Gao and Yang, 2023) or ultimately promote skepticism about the review (Kong and Teng, 2023), which affects their comprehension speed and feedback acceptance (Mahfoodh, 2017). Different types of feedback lead to different emotional experiences for learners. For example, some feedback content is highly critical, and the expression is direct and face-threatening, which is likely to evoke strong negative emotions (Madhu and Hu, 2021). Some researchers believe that feedback can easily arouse negative emotions and thereby undermine learners' enthusiasm (Truscott, 1996). However, the study of emotions in L2 academic writing from the perspective of positive psychology takes a holistic view of the integration of positive and negative emotions (MacIntyre et al., 2019) and has reached different research conclusions. For example, some learners experience gratitude,

disappointment, happiness, embarrassment (Han and Xu, 2020), frustration (Zheng and Yu, 2018), relief, and excitement after receiving feedback (Han and Hyland, 2015; Mahfoodh, 2017), revealing the complexity and situational and dynamic nature of emotions in L2 writing. The study by Han and Hyland (2019) was among the earlier studies exploring emotional experiences in L2 feedback, and used Pekrun (2006) and Pekrun and Linnenbrink-Garcia (2012) emotional analysis framework to explore English second language learners' emotional experiences in feedback, providing references and examples for subsequent research on L2 learners' emotional experiences elicited by feedback. The study used qualitative research methods supplemented by classroom observation recordings and students' facial expressions for triangulation, revealing two learners' emotional experiences, such as calmness, guilt, and nervousness, in response to written corrective feedback (Han and Hyland, 2019). Geng and Shulin (2024) study showed that learners had 65 discrete emotions after receiving feedback, the most common of which were negative emotions, but these negative emotions did not discourage them because they actively sought external resources to improve their revisions. It can be seen that the previous view that feedback always elicits negative emotions may not be true (Truscott, 1996; McMartin-Miller, 2014) and perhaps does not reflect the actual experience of second language learners.

Some studies have explored learners' emotion regulation after receiving feedback and found that learners can promote positive emotions through a variety of strategies (Liu and Yu, 2022). Teachers can also use psychological interventions to create a positive language

learning environment (MacIntyre et al., 2019) and guide learners to develop learning emotional intelligence. Most studies are based on large-scale quantitative analyses. For example, Han et al. (2024) conducted a questionnaire survey on 363 postgraduate students and found that they can adopt strategies such as emotion-oriented regulation and situation-oriented regulation to regulate emotions; Teng and Ma (2024) demonstrated the importance of self-perceived motivation and confidence in regulating students' emotions and proposed that these findings should be used throughout the feedback process to facilitate students' seeking, generating, processing, and using feedback. Through focused writing and semi-structured interviews with middle school students, Gu et al. (2022) found that seeking social interaction, developing competence, and cognitive reappraisal were the most commonly used strategies by students. At the same time, specific academic emotion regulation strategies are related to the purpose of emotion regulation. For example, students most often use the strategy of developing competence when they want to learn from their peers' essays.

In general, current research on L2 writing feedback focuses on its role and different methods. The little attention paid to learners' emotions mostly revolves around the changes in emotional experiences before and after a particular feedback (Ellis, 2010) and has not extended the observation to dynamic situations such as longer periods of time and multiple writing sessions. Furthermore, academic emotions and emotion regulation have generally been studied in English L2 writing contexts or in the context of master's or Ph.D. theses (e.g., Mirka and Kirsi, 2019; Geng and Yu 2024), with little attention paid to Chinese L2 undergraduate students. As these learners are new to academic writing and studying in a foreign culture, their emotional experiences are richer and more complex than those of learners who study L2 in their home countries and those who have more writing experience, and thus merit in-depth exploration through qualitative means such as case studies, as they can offer unique insights into the complex interplay between emotions and learning in a cross-cultural educational environment.

In addition, previous studies have found that an individual's emotional experience is not only influenced by the feedback itself but also generated by the interaction of many factors, such as individual and social context (Jin and Zhang, 2018; Jiang and Dewaele, 2019). Regrettably, no research to date has delved into the specific contextual factors that shape these emotional experiences. To enhance our understanding of how CFL students navigate their emotions and employ emotion regulation strategies in response to feedback and to extract implications with far-reaching instructional significance and applicability across diverse educational settings, the current study endeavored to address the following two research questions:

1. How do CFL students' emotional experiences change throughout the teacher's three feedback sessions and each time's revision process?
2. What regulation strategies did CFL students use to self-regulate the academic emotions aroused by the teacher's feedback?

Research methods

Research context

The data for this study come from a 16-week academic Chinese writing course for international undergraduates offered by a Chinese university in the spring semester of 2023. The course was taught in

Chinese and aimed to cultivate students' awareness of genre and norms in academic writing, which could prepare them for future academic research and dissertation writing. There are three summary writing tasks in one semester (in the third, eighth, and twelfth weeks, respectively). The students are required to screen and extract the key points of the original text based on reading and understanding an academic paper and write a summary in their own language. The instructor is a native Chinese-speaking teacher from China, Audrey (all names are pseudonyms), who majored in Chinese language and literature in her undergraduate, master's, and doctoral degrees and has 25 years of experience in teaching Chinese L2 academic writing. Before assigning the writing task, the teacher (Audrey) had already taught the methods of summary writing in class. The students completed the writing in class, and the teacher (Audrey) gave each student written feedback and oral feedback in class.

Participants

This study aimed to select students so that there would be individual differences among the participants in terms of writing experiences and emotional experiences. We used maximum purposive sampling, and two students, Yuki and Sala, agreed to participate. The participants are from different countries around the world. Due to the uneven development of Chinese language proficiency, writing experiences, and beliefs and attitudes toward academic writing feedback, their emotional experiences of academic writing feedback are also different. As the researchers are involved in teaching the course, they are very familiar with the participants. We recruited the participants through the first researcher, and they will provide information about their feedback and academic writing experiences. The participants were purposively selected based on three specific criteria: (a) they both had a strong desire to participate in the research and were willing to share their experiences and emotions in academic writing, which provided feasibility for the smooth progress of the research; (b) they both needed to write their dissertations in Chinese when they graduated in the future; (c) they had different HSK (Chinese Proficiency Test) level 6 writing scores, previous academic writing performance, first summary writing scores (Yuki was a high-level writer and Sala was a low-level writer), the feedback they received, and their academic emotions showed an undeniable contrast. The two cases can reflect the students' emotional experiences and regulation in writing feedback. Yuki's and Sala's backgrounds are given in Table 3.

It is worth noting that only two learners were selected as subjects in this study. The purpose of doing so was to choose individuals with specific characteristics or experiences and thoroughly explore their detailed experiences and reactions in specific situations. This helps to focus on the research questions and avoid the situation where the information becomes too scattered due to an overly large sample, making it impossible to deeply analyze the relationships between key factors. In addition, diverse data collection methods were used for these two students in this study to obtain comprehensive and in-depth data, thus compensating for the deficiency of the small sample size.

Data collection

The emotional reactions to the teacher's feedback of student participants were mainly investigated through interviews,

TABLE 3 The CFL students' background information.

	Yuki	Sala
Gender	Female	Female
First language	Japanese	Thai
Major	Journalism and Communication	Journalism and Communication
Year of college	First-year	First-year
Writing score of HSK-6	75	60
Summary writing score 1	84	79
Summary writing score 2	87	80
Summary writing score 3	89	88

retrospective oral reports, audio recordings, and class observation notes, and the teacher's written feedback was also collected as [Supplementary Data](#). That is, all the methods mentioned above are used to answer Question 1. The emotion regulation strategies were analyzed mainly through interviews and retrospective oral reports. Thus, the two methods are used to answer Question 2.

- Interviews: Three formal, semi-structured interviews (see [Supplementary Appendix 1](#) for the interview guidelines) and multiple informal interviews. The first interview was conducted at the beginning of the semester and lasted approximately 40 min, mainly to explore learners' personal information, experiences of academic writing and dealing with feedback, and beliefs about academic writing and feedback. The next two interviews focused mainly on learners' emotional experiences of each piece of writing feedback and future academic writing and regulation strategies. In the opening part, the first researcher explained the purpose and procedure of the study. In the questioning part, the participants were encouraged to give detailed insights into their emotional reactions to the feedback and how they regulated their emotions to proceed with the writing. The semi-structured interviews were conducted within 48 h of the learners receiving the first (F1), second (F2), and third (F3) feedback, respectively, and each interview lasted approximately 60–90 min. The informal interviews were mainly based on daily communication via WeChat and face-to-face communication, with the aim of supplementing the information not obtained in the formal interviews and tracking the real state and inner experience of the participants in their learning life.
- Retrospective oral report: students conducted a 5–10 min retrospective oral report within 24 h after receiving the three feedback. Students were asked to report on their emotional experiences and regulation strategies before and after receiving feedback, and the researcher recorded their responses to the feedback. To stimulate recall, participants were able to review the written text, the teacher's written feedback, and the error log to recall their experiences and attitudes toward feedback.
- Class observation notes: One of the researchers was a teaching assistant for the Chinese academic writing course, who was able to observe the class in depth and know the language level and writing ability of the participants.
- The teacher's written feedback: Audrey gave feedback and scores to each participant's summary writing text. This data can

reflect the participants' writing ability and the type of feedback given by the teacher.

We also took the following measures to keep the participants' information confidential: During data collection, each participant was anonymized, and pseudonyms were used to label all content, such as interview records and reflection logs; in the data processing stage, the data was stored in password-protected software, and only the researcher could directly access the original data; in the data use stage, the research members signed a confidentiality agreement, promising not to disclose the data content; in the paper writing stage, the identity information of the participants was thoroughly anonymized to ensure that it could not be traced back to specific individuals.

In addition, in order to prevent adverse emotional reactions of the subjects during interviews and self-reports, we also took a series of relevant measures. First, when recruiting the subjects, we informed them of the possible emotional challenges involved in the research process, such as the adverse emotions that might be triggered by recalling difficult experiences in the learning process and facing critical feedback. We ensured that the participants voluntarily participated in the research with full knowledge and provided them with the right to withdraw at any time. Second, the researcher (the first author of this article) familiarized himself/herself in advance with the types of possible adverse emotional reactions and the corresponding coping methods and provided the subjects with information about the school's mental health center for them to seek help. Third, during the research activities, such as interviews and classroom observations, we closely monitored the emotional states of the participants. Once we found that the participants had emotional fluctuations, we immediately suspended the research activities and, at the same time, gave the participants the opportunity to express their emotions. We used active listening and encouragement techniques. Finally, we conducted follow-up visits to the participants to understand their emotional recovery situations and reminded them that they could continue to seek help if they still had emotional problems.

Data analysis

Analysis of the teacher's written feedback

Writing text analysis is mainly concerned with defining the amount of feedback given by the teacher. According to [Hyland \(1998\)](#), a single feedback point is defined as each opinion expressed on a particular aspect of the article, including each comment and revision.

This study refers to the classification of Ellis (2009) to classify feedback points from two dimensions: one is non-corrective feedback, which refers to the teacher's comments on the writing (Ferris, 1997); the other is corrective feedback, including direct feedback (correcting errors directly) and metalinguistic feedback (explaining the metalinguistic rules violated by the error). The specific content is shown in Table 4.

Analysis of other data

Data analysis adopts the qualitative content analysis method (Miles et al., 2013; Han and Xu, 2020). Individual case files are established for the data, and qualitative analysis is carried out according to the steps from within - individual cases to cross-individual cases (Han and Hyland, 2019). Data were reviewed by thematic analysis and the constant comparative method of analysis, in which the data were systematically analyzed through a three-stage process of first and second-cycle coding for data condensation (Miles et al., 2020). The coding process was as follows: First, the within-case analysis stage. The recording was transcribed immediately after each interview, and data analysis began immediately after review and approval by the participants. Micro-analysis of the data was carried out after the first written feedback from the first participant. The data were read repeatedly. Using the academic emotion classification framework (Pekrun, 2006; Pekrun and Linnenbrink-Garcia, 2012; Han and Hyland, 2019) and the emotion regulation strategy classification framework (Pekrun, 2006), relevant data fragments related to the research questions were cyclically coded and classified, and the analysis framework was adjusted according to the data. Emotional experiences and regulation strategies were mainly identified using vocabulary from student interviews, self-reports, and classroom observation notes. Then, a preliminary coding list was produced, including emotional

experiences and regulation strategies. Second, the data of the second participant were analyzed, and the coding list was improved. Thirdly, in the cross-case horizontal comparison stage, case narratives are completed by refining themes and taking an inductive approach. The coding lists of the two participants were compared to select, organize, and merge important concepts and dimensions. For example, 'immediate emotions' and 'emotions after reading feedback' were grouped into one category and coded as 'emotions during academic work.' The frameworks (Pekrun, 2006; Pekrun and Linnenbrink-Garcia, 2012; Han and Hyland, 2019) were adapted according to the specific content of the data:

(a) Five discrete emotions were added to Pekrun and Linnenbrink-Garcia (2012) taxonomy: conflict, the feeling of being torn because you do not know exactly which way is better for you; novelty, the feeling of being strange and new because of something you have never seen or experienced before; achievability, the feeling when desires and reality are in balance; relief, the positive feeling induced by the fact that the expected failure did not actually occur; anticipation, the feeling of longing and yearning for something unknown in the future. Definitions of all discrete emotions are given in Supplementary Appendix 2. (b) Through a detailed review of the interview, oral report, and classroom observation of the present study. It was found that when facing academic writing feedback, the participants more often adopted other types of strategies (such as emotion-oriented regulation, appraisal-oriented regulation, and situation-oriented regulation strategies) to deal with emotional problems and did not clearly exhibit the behavior pattern of directly solving the problem itself (i.e., "problem-oriented strategy"), such as seeking additional learning resources to specifically solve the problems exposed in writing or changing the learning method to avoid similar problems from recurring. Therefore, it was removed from the classifications.

TABLE 4 The CFL students' feedback information.

Name		Non-corrective feedback (comments)	The teacher's written feedback
Yuki	F1	Ideological content comments (negative) 3 Overall improvement suggestions 1	Direct feedback 9 Indirect feedback 0 Metalinguistic feedback 3
	F2	Organizational structure comments (positive) 1 Language usage (negative) 1	Direct feedback 10 Indirect feedback 0 Metalinguistic feedback 1
	F3	Language usage comments (positive) 1 Ideological content comments (positive) 1 Overall improvement suggestions 1	Direct feedback 9 Indirect feedback 0 Metalinguistic feedback 2
Sala	F1	Ideological content comments (positive) 1 Overall improvement suggestions 1	Direct feedback 15 Indirect feedback 0 Metalinguistic feedback 2
	F2	Ideological content comments (positive) 1 Organizational structure comments (negative) 1 Overall improvement suggestions 1	Direct feedback 15 Indirect feedback 0 Metalinguistic feedback 6
	F3	Language usage comments (negative) 1 Ideological content comments (positive) 1 Organizational structure comments (negative) 1 Overall improvement suggestions 1	Direct feedback 16 Indirect feedback 0 Metalinguistic feedback 5

Except for identifying discrete emotions, informed by research on academic emotions, a dimensional approach was also taken to identify the object focus, valence, and activation of emotions. The discrete emotions were categorized by object focus following Pekrun and Linnenbrink-Garcia (2012) classifications (see Table 1). This study did not find that the feedback aroused the learners’ “topic emotions” (emotions related to the content of the writing task). Therefore, it was removed from the analytical framework. Pekrun (2006) assigns different valences and activations to discrete emotions, e.g., ‘anxiety’ belongs to the negative-activating dimension and ‘excitement’ to the positive-activating dimension. However, based on the interviews and participants’ reports, we found that the degree of activation of the same emotion varied. For example, Yuki felt anxious before receiving the three feedback sessions, but she thought that the activation of anxiety at the first feedback was the strongest and then gradually decreased. Therefore, in this study, the activation of these three anxiety emotions was coded as activating, neural, and deactivating, respectively.

In addition, to avoid the influence of the participants’ L2 proficiency on the expression of meaning, this study used a triangulation method for the assessment of emotions and strategies. Except for the analysis of interviews and self-reports, the following data were also included: phonetic details and gestures during interviews, self-reports, students’ reflection logs, and the researcher’s classroom observations. Finally, the above data were archived by case, followed by horizontal comparisons across cases, and the case narratives were completed by refining themes. Two researchers

independently coded the above data and initially achieved 83% inter-coder agreement. The researchers then discussed the different codes of discrete emotions, valence, activation, and emotion regulation in order to resolve the differences. Inter-coder reliability eventually reached 96%. The coding schemes for CFL students’ emotional reactions to feedback are shown in Table 5, and the academic emotions categorized by object focus and examples are shown in Table 6.

Findings

Yuki: academic emotions and regulation strategies under self-imposed pressure

Yuki comes from a bilingual Chinese–Japanese family in Japan. She attended high school in northeast China. She was very concerned about her GPA and planned to study for a master’s degree, but she did not know how to express herself in an academic genre. Yuki believed that summary writing and feedback could improve academic writing skills, but the skills gained from feedback on a specific piece of writing could not be transferred to other papers. Yuki’s overall emotions aroused by feedback were “strongly changing and up and down” (the third oral report) and constantly evolving between negative and positive, even though her scores on three pieces of writing were among the best in the class. Based on the oral report and the interviews, the researcher summarized the key times of emotional change and the changes in Yuki’s academic emotions (see Table 7).

TABLE 5 Coding schemes of emotional responses to feedback.

Codes of emotional responses to feedback	
Discrete emotions	Conflict, anxiety, hopelessness, guilt, novelty, confusion, tranquility, achievability, gladness, trust, hope, expectancy, gratitude, relief, curiosity, and satisfaction
Object focus	Achievement emotions (including prospective outcome emotions, retrospective outcome emotions, and activity emotions), epistemic emotions, and social emotions
Valence	Positive, neutral, and negative
Activation	Activating, neutral, and deactivating

TABLE 6 Academic emotions categorized by object focus and examples in feedback situations.

Category	Discrete emotions	Examples
Achievement emotions		
-Activity emotions	Novelty, confusion, anxiety, gladness, relief, achievability, tranquility	Yuki: Maybe I am a bit unsophisticated. It was the first time I saw the revisions on the computer (oral report) Sala: I was very happy after reading the comments, and I found that the teacher praised me (interview)
-Prospective outcome emotions	Conflict, novelty, expectancy, anxiety, tranquility, hope	Yuki: Before receiving the feedback, I felt conflicted (reflection log); Sala: This time should be an improvement over the last time. I was quite tranquil (reflection log)
-Retrospective outcome emotions	Confusion, guilt, tranquility	Yuki: I feel guilty, as always, for wasting the teacher’s time (oral report)
Cognitive emotions	Hopelessness, guilt, curiosity	Yuki: I have tried my best, but it seems that my ability is still not enough (interview)
Social emotions	Tranquility, gratitude, anxiety, trust, relief	Yuki: I am grateful to the teacher. She spent a lot of time revising (reflection log) Sala: I am worried that my Chinese is not good enough. I am really anxious to meet the teacher (oral report)

TABLE 7 Yuki's academic emotions to three feedback sessions and regulation strategies.

		Before receiving	Feedback receiving	Feedback reading	Oral feedback in class	After revision
F1	Academic emotions	Conflict anxiety	Hopelessness guilt	Novelty confusion anxiety	Tranquility gratitude	Confusion guilt
	Valence-Activation	Negative-activating Negative-activating	Negative-activating Negative-activating	Positive-neutral Negative-deactivating Negative-activating	Neutral-neutral Positive-activating	Negative-Deactivating Negative-activating
	Strategies	Emotion-oriented regulation	Situation-oriented regulation			
F2	Academic emotions	Anxiety	Relief achievability	Confusion	Glad trust	Tranquility
	Valence-Activation	Negative-neutral	Positive-neutral Positive-neutral	Negative-deactivating	Positive-neutral Positive-neutral	Neutral-neutral
F3	Academic emotions	Anxiety	Tranquility	Confusion	Glad	Hope
	Valence-Activation	Negative-deactivating	Neutral-activating	Negative-neutral	Positive-neutral	Positive-neutral

A rough start: walking alone in anxiety and guilt in the first feedback session

The most common emotions Yuki experienced when she first received feedback on her academic writing were guilt and anxiety. After finishing her first piece of writing, Yuki felt anxious and conflicted as she waited for the feedback email. “I’m always worried about the grade. I hope the teacher will revise it more so that I can learn a lot, but the score will be low” (first oral report). So, she quickly looked at the score as soon as she received the email.

[First interview].

“My first reaction was, ‘It’s over.’ I did not even get 85. I tried my best, but it seems that I’m still not good enough. I may have to catch up on this GPA in another course. Then came that familiar feeling of guilt.”

During the interview, she mentioned the emotions of “hopelessness and guilt,” even though it had been 36 h since she had received the feedback, and her eyes were red, showing that she was in a negative-activating state. After looking at the score, Yuki prepared to read the feedback carefully. What caught her eye was the revision that Audrey had made using the ‘word revision mode.’ It felt very novel. However, much of the content of the feedback made her feel confused and anxious:

[First interview].

“I’m not used to the ‘Word revision mode.’ The teacher also said that a sentence was repeated, but I did not really notice it. To be honest, I do not know if it was because I was too anxious and could not think, which led to me not being able to understand these things after they were revised.”

When talking about emotion regulation that could be used to improve these emotions, Yuki stated that she “refused to regulate” and that “living in anxiety is especially good because it can motivate me to study harder. Maybe only grades can make me really happy” (oral reports).

The teacher then gave oral feedback to the class on the common writing problems, which resolved most of Yuki’s confusion. Audrey praised the students’ efforts, smiled, and nodded to encourage everyone. Yuki also had social emotions such as gratitude. “If the teacher did not tell me some problems, I might never know them in my life” (first interview). Yuki aroused social emotions (gratitude) to “hedge” negative emotions through emotion-oriented regulation strategies, which diverted her attention from the negative emotions and made her more diligent and engaged (Han and Xu, 2020).

After the revision, Yuki recalled that the feedback on this writing was still negative, and she felt very guilty. She was still confused about the content of the feedback: “I do not know if I can still write questions when I write papers in the future.” (Oral report) But Yuki said that she would not ask the teacher for advice during class breaks for fear of wasting the teacher’s time and the teacher having a bad impression of her (oral report and class observation), even though feedback is a process in which learners should actively seek, generate, process, and use feedback to apply new knowledge in current or subsequent writing tasks (Teng and Ma, 2024). What is more, growing up in the East Asian cultural circle made her not want her classmates to see her competitive spirit. Feeling guilty, Yuki said that there was no effective way to adjust, and she could only use situation-oriented strategies to avoid feedback sessions that led to her negative emotions. “I can only say do not think about it for now, put it aside, and look at it later.” (Oral report).

An adaptive state: groping in the collision of positive and negative emotions in the second and third feedback sessions

After the first feedback session, Yuki had a preliminary understanding of the basic methods of summary writing and had reasonable expectations about academic outcomes and possible writing problems (the third oral report). Therefore, the activation of anxiety arousal before and after the second and third feedback sessions gradually deactivated.

[Second interview].

“I still look at the results after the feedback: 87. The score is higher than last time and also above 85. I feel relieved. The lower the score, the more upset I will be.”

Grades, the result of the writing activity, and the extrinsic value of the writing (Pekrun, 2006) were the most emotionally evocative feedback for Yuki. 85 was the watershed of the normal distribution of grades in her university and also the watershed of her positive and negative emotions. For grades that improved and were above 85, she presented the academic activity emotion: tranquility. Yuki also received positive comments in the second and third feedback sessions, which made her believe that she could make progress as long as she worked hard (second interview). It could be said that grades can pull her from the bottom of the negative-activating emotional experience to the positive-neutral emotional experience, but the teacher's positive and friendly encouragement can soar her emotions to the positive-activating state, which has a powerful reshaping effect on Yuki's confidence.

However, there was still confusion in the two writing feedback sessions, and the feedback in class was still good medicine to remove the confusion: it was replaced by emotions such as trust in the teacher or gladness to have "another example to refer to for the next writing." "The teacher's encouragement and attitude gave me the motivation to keep working hard" (second interview). As a result, Yuki focused on revising and gradually felt more confident about academic writing. By the third writing, she had basically mastered the writing skills and could estimate the gains and losses of the scores in writing before receiving the feedback (third oral report). In the interview after the last feedback, she used "a little" to describe all her emotional experiences, e.g., "a little anxious." The accumulation of writing knowledge and skills led Yuki to be full of hope for future writing.

Sala: academic emotions and regulation strategies in happy growth

Sala is from Thailand. She started learning Chinese because her idol was a Chinese singer. Sala planned to work in a media company in China and write copy to promote the idol's career, so she put a lot of emphasis on her writing skills. Sala did not care much about her academic grades, believing that as long as she learned and did her best in every course and assignment, she would be fine. She looked forward to receiving feedback on her writing and strongly believed

that feedback could help her improve her writing logic and language. Compared to Yuki, Sala's Chinese writing level was lower, but she made three times more progress in her writing grades, and her emotions were relatively stable. The key times of emotional change and the changes in Sala's academic emotions are shown in Table 8.

A steady start: acquiring knowledge in tranquility in the first feedback session

Sala's overall assessment of her emotions was "very stable, not particularly happy or unhappy" (third oral report), and the analysis results also showed that this was indeed her emotional experience during the three feedback sessions. After the first summary writing, Sala said that the teacher's feedback was more important to her than the score. She valued her writing level and academic outcomes and had only a little anxiety before receiving it.

[First oral report].

"I am a bit anxious about the score. I tried my best, and any score was fine. 70–80 is my level of writing. Above all, I want to know what I need to improve."

After receiving the feedback, Sala felt that "it was fine, I did not feel happy or unhappy, the score of 79 was in line with what I expected of myself" (first interview). The academic results that were in line with expectations made Sala feel calm. Later, when reading the feedback, Sala, like Yuki, had difficulty reading because of the revision mode of the document, which caused confusion. She cautiously read the specific content of the feedback and gradually got used to the way of looking at the feedback:

[First interview].

"I found that I had grasped the wrong key points, and some expressions were too abstract, and the language was not academic enough. I paid more attention to how I could improve my writing skills, so I was very pleased to see so many revisions."

TABLE 8 Sala's academic emotions to the three feedback sessions and regulation strategies.

		Before receiving	Feedback receiving	Feedback reading	Oral feedback in class	After revision
F1	Academic emotions	Anxiety Expectancy	Tranquility	Confusion expectancy	Satisfaction	Anxiety Relief
	Valence-activation	Negative-deactivating Positive-neutral	Neutral-neutral	Negative-neutral Positive-activating	Positive-activating	Negative-deactivating Positive-neutral
	Strategies	Emotion-oriented regulation				
F2	Academic emotions	Anxiety	Glad	Tranquility	Trust	Tranquility
	Valence-activation	Negative-deactivating	Positive-activating	Neutral-neutral	Positive-activating	Neutral-neutral
F3	Academic emotions	Tranquility	Glad	Curiosity	Tranquility	Hope
	Valence-activation	Neutral-neutral	Positive-activating	Positive-activating	Neutral-neutral	Positive-activating
	Strategies	Appraisal-oriented regulation				

Acquiring new writing methods from feedback was the most meaningful and valuable thing for Sala. She was satisfied with the new knowledge she had learned, although the confusion still existed. During the interview, Sala's arms were relaxed and naturally placed on her legs. Her high-pitched voice and constant laughter reflected her positive emotions (Han and Hyland, 2019). Later, the confusion was resolved during the feedback session in class. Sala was looking forward to using new methods in her future writing:

[First interview].

"I did not know that breaking it down into smaller paragraphs would make it clearer until the teacher explained it to me in class. Wow, I can write like that next time."

Sala then devoted herself to revision and asked Audrey for advice after class on the new confusion. However, the one-on-one oral feedback session triggered Sala's complex emotional experiences: at first, she was worried about the poor communication with the teacher and felt slightly anxious (she spoke faster and covered her mouth to chuckle when she talked about this), then she began to watch the idol's inspirational videos to encourage herself and made up her mind to ask the teacher for advice. The subsequent communication with the teacher aroused Sala's relief and prevented her from being hit: "Even if I blurted out two words, she could understand what I was talking about... I was much clearer after asking her" (first oral report). Moreover, the feedback session in class improved Sala's self-efficacy and motivated her to invest in the learning process. Sala used an emotion-oriented strategy, that is, watching videos to motivate herself and distract her feelings of anxiety, and finally achieved successful revision (Oxford, 2016).

A joyful state: growing rapidly under the infiltration of positive emotions in the second and third feedback sessions

Before receiving F2 and F3, Sala had already noticed her progress in writing while writing, so her emotions gradually became tranquil. She began to receive positive comments: the structure was very clear, and the language was in line with the norm (second feedback text). "I was very glad after reading the comments, and I found that the teacher praised me!" (oral report). Sala's writing scores increased in both the second and third feedback sessions; the scores were 80 and 88, respectively. However, she believes that the teacher's comments have the greatest impact on her emotions, much more so than the score itself. Clearly, the content of positive feedback increases students' self-regulated learning and confidence in the learning task (Mirka and Kirsi, 2019), which in turn generates more positive emotions (Pekrun, 2006).

F3 also evoked Sala's cognitive emotion of curiosity: "The teacher marked that this sentence was wrong. I also looked through my other writings to see if there was the same mistake, and sure enough, there was" (third interview). The mistakes marked by Audrey reminded Sala of similar problems in previous papers. Feedback can help her transfer language knowledge to multiple studies.

F2 in class still aroused Sala's trust in the teacher. She believed that it could effectively sort out writing errors and consolidate writing

knowledge based on written feedback. Sala gradually adapted to and got used to the revision, and her emotional experience became tranquil. After the last revision, according to the feedback, Sala found that she was more familiar with academic writing and had made significant progress (class observation). At the same time, Sala found that she had become more manageable with writing in other courses: "There was a class that asked us to write a book report. I introduced the book in four points (research background, research questions, research methods, and evaluation) according to the teacher's feedback, and the quality was very good" (second interview and oral report). The use of evaluative strategies made Sala feel hopeful about the feedback.

Discussion

Informed by the analytical framework of academic emotions (Pekrun and Linnenbrink-Garcia, 2012) and emotions regulation strategies (Pekrun, 2006), this study examined the emotional experiences and emotional regulations of two Chinese academic writing novices in the context of three writing feedback sessions in one semester. The data showed that before and after receiving feedback, the two CFL students experienced as many as 16 kinds of achievement emotions, cognitive emotions, and social emotions with different levels of valence and arousal, and they could use a variety of strategies to regulate emotions. These emotional experiences were intertwined, fluctuating, and evolving along the timeline of "before receiving feedback, just after receiving feedback, and during in-class feedback in 1 to 5 days after revision" (Han and Xu, 2020). These changes occurred within 1–5 days after each feedback session and throughout the semester across three feedback cycles.

The above research results further confirm the framework proposed by Pekrun (2006) and Pekrun and Linnenbrink-Garcia (2012) for analyzing academic emotions and their regulatory strategies. In addition, by incorporating emotions such as conflict, novelty, achievability, relief, and anticipation, the study suggests that the taxonomy of academic emotions (Pekrun and Linnenbrink-Garcia, 2012) needs to be expanded and adapted for feedback sessions. The academic emotions aroused by feedback underwent complex changes, not merely shifting from negative to positive. Emotions could be transient or persist throughout the entire feedback processing cycle (Han and Xu, 2020; Kikuchi and Lake, 2021). This result differs from previous studies, which state that feedback always triggers negative emotions (Truscott, 1996). However, the emotions of the two CFL students gradually became positive in the three feedback sessions. Even Yuki, who often experienced anxiety at the beginning, gradually felt tranquil and grateful in the second and third feedback sessions. The same discrete emotion varied along the activation. For example, before receiving feedback, both students felt anxiety, but the activation was different: Yuki was activating, while Sala was only deactivating. This finding forced us to adjust the taxonomy of emotions as seen in previous studies (Pekrun and Linnenbrink-Garcia, 2012). Emotional reactions with the same valence but different activation levels had an impact on students' motivation and learning effects (Pekrun et al., 2002; Geng and Yu, 2024). Yuki was trapped in activating anxiety and had difficulty engaging in

the subsequent modification after receiving feedback. At the same time, Sala quickly entered the learning state and immediately took corrective actions, indicating that learners with stronger psychological resilience will actively deal with negative emotions, think about the problems presented in the feedback, and seek value from it (Tugade and Fredrickson, 2004).

We found that compared to written feedback, oral feedback in class was always a good medicine to relieve their nervousness and anxiety, which can help them understand the content of feedback and regulate their emotions. The results also showed that the discrete emotions aroused by the two learners at different stages of feedback showed similar tendencies. Written feedback from the teacher may be more likely to evoke negative emotions in learners. In the learners' process of dealing with written feedback, firstly, in terms of prospective outcome emotions, both students exhibited negative emotions before receiving feedback and felt nervous and anxious about the score and the content of the feedback because feedback is usually regarded by learners as an evaluation of the quality of academic writing (Mahfoodh, 2017; Geng and Yu, 2024). Second, in terms of activity emotions, both learners had different negative emotions, such as anxiety and confusion, in the first feedback (F1) session when they were engaged in the feedback activities (receiving and reading the feedback); fortunately, these negative emotions did not affect their revision and reflection on their writing too much, and their emotions gradually tended to be neutral and positive in the following two feedback sessions. Finally, after revising their writing according to the feedback, both of them showed retrospective outcome emotions such as tranquility and hope as the number of feedback sessions increased. It can be seen that engagement with feedback and new knowledge may be able to smooth out learners' negative emotions to a great extent and increase their happiness and satisfaction with learning.

In addition to academic achievement emotions, Yuki and Sala also experienced epistemic emotions such as novelty or confusion when they first encountered feedback on Word's "document revision mode." They regarded the feedback as a manifestation of the teacher's concern for their learning and the teacher's serious work (Lee, 2008a,b; Han and Xu, 2020), so the feedback also awakened their social emotions, such as gratitude and relief. These social interactions with the stakeholders also helped novice researchers enter the academic discourse community (Geng and Yu, 2024).

With the increase in the frequency of feedback sessions, the emotions of students gradually tended toward positivity. In addition to their gradual adaptation to the feedback content and improvement in writing skills, they also benefited from the appropriate use of their emotion regulation strategies. Both learners used emotion-oriented strategies. For example, Sala adopted an emotion-oriented regulation strategy of watching idol videos repeatedly to relax and reduce negative emotions when faced with anxiety during the one-on-one feedback session with the teacher. The idol was both a motivation for her to learn Chinese and could encourage her to bravely ask the teacher for advice. It can be seen that positive emotions can reduce negative emotions to a certain extent and help learners reduce the destructiveness of negative emotions (Oxford, 2016; Han and Xu, 2020; Gu et al., 2022). However, when Yuki, who used the situation-oriented strategy, faced feedback that put her in a

negative mood, she chose to avoid the feedback text that triggered her negative emotions. As she grew up in the implicit and introverted culture of East Asia, she was concerned about revealing her true self to her classmates. Although she still had doubts and confusion about the feedback and was under great pressure from her GPA throughout the year, she was not encouraged to seek support and solve problems from teachers or peers. In comparison, the appraisal-oriented strategy was more effective in regulating emotions and had a far-reaching impact. Sala used the appraisal-oriented strategy, and her positive attitude toward the role of feedback (positive value appraisal) and confidence in her progress (positive control appraisal) led her to pay more attention to feedback (Pekrun, 2006) and increased her satisfaction with revision. Finally, she was full of hope for future academic writing.

In addition, previous studies have suggested that learners with low language proficiency are more likely to receive more feedback and, therefore, may experience more negative emotions with a higher degree of activation (Hyland, 1998). However, this study found that this hypothesis may not reflect the actual experience of CFL students: Yuki, whose language proficiency and writing scores have been among the top all along, has mainly negative emotional experiences, while Sala, whose language proficiency and scores were initially low, is mostly in a positive mood. This finding suggests that L2 learners' emotional experience in feedback is not solely influenced by language level and feedback content. To explore the factors that affect learners' emotional experience, it is necessary to combine a wider range of situational factors (Bruton, 2009, 2010). First, different forms of teacher feedback have different effects on learners' emotions. Written feedback is usually more likely to arouse learners' negative emotions. Compared with the single-modal (visual) feedback input in writing, the auditory and visual multimodal feedback of teachers' tone, expression, and actions in oral feedback were more helpful in solving learners' confusion and alleviating their negative emotions. After receiving one-on-one feedback or feedback in class, the two CFL students mostly turned to positive emotions, which proved to be a good corrective effect of oral feedback.

Second, the students' perceptions of the controllability of writing outcomes could also evoke their different emotional experiences. For example, the lower score in the first feedback session made Yuki think that her writing was a failure, thus lowering her subjective control assessment of the writing task, and negative emotions were generated as a result. The improvement of her writing scores and the appearance of positive comments in the following F2 and F3 helped her realize her ability to control the outcomes of the academic activities, and her emotions gradually became positive. Finally, this study also found that paying attention to the internal value of feedback could help the students experience more positive academic emotions. Both CFL students had a high level of engagement, while their subjective value appraisal of feedback was significantly different: Sala attached importance to internal value and believed that revising feedback could acquire writing knowledge. She often had emotional experiences of gladness and achievement from mastering writing skills. Therefore, she paid more attention to feedback regardless of whether she received good grades or not. However, Yuki's high investment in revising feedback was mostly driven by the desire for a high GPA (external value) and viewed feedback itself as a tool to achieve academic

achievement (Papi, 2010). Therefore, when academic outcomes that did not meet expectations appeared, Yuki could only feel a lot of negative emotions, regardless of whether she had mastered writing skills during the revision feedback process.

The findings further support the view that emotion research from the perspective of positive psychology needs to be contextualized (Han and Xu, 2020). Judging from the academic achievements of the three summary writing tasks, both Yuki and Sala, the high-level and low-level Chinese writers, respectively, undoubtedly made progress. However, Sala, who “did not care much about the scores,” made more remarkable progress. Yuki kept her eyes fixed on the goal and just kept going and going. Therefore, it is no wonder that the crazy attacks of tiredness and anxiety came. Guided by the goal of pursuing scores and GPA, Yuki actively conducts self-discipline and self-management and gradually grows into a “calculable person” (Foucault, 2012). The findings of this study support the control-value theory (Pekrun, 2006). Moderate self-imposed pressure is a condition that can promote deep learning. However, excessive self-imposed pressure may heighten docility, devalue creativity, and thus hinder the occurrence of real learning (Lin and Lin, 2023). At this time, knowledge is overshadowed by the practical, insignificant, and dull, and the time for students to develop critical thinking and exploratory thinking is immediately compressed indefinitely. Eventually, the emphasis on the learning style for the purpose of assessment and the excessive self-imposed harsh requirements for emotions lead to a decline in learners’ motivation and morale and induce learners’ hopelessness and anxiety toward academic and learning activities in terms of emotions (Lin and Lin, 2023).

Sala was eager to write knowledge and enjoyed the learning process. She not only has a positive value appraisal of summary writing and its feedback but also has sufficient confidence to make progress with the help of teachers’ feedback and looks forward to positive future academic achievements. She walked, enjoyed the flowers, and grew strongly accompanied by happiness and tranquility. The accumulation of state emotions that occur frequently will form trait emotions (Pekrun et al., 2011). If students can effectively regulate their emotions and often experience positive academic emotions, they are more likely to build positive personality traits, thereby gaining more positive psychological resources to cope with challenges and ultimately achieve a double harvest of academic achievement and happiness (Han and Xu, 2020).

Conclusion

This study provides empirical data on how to improve happiness in the process of learning Chinese second language academic writing from the perspective of positive psychology. Compared with previous studies on feedback texts and feedback emotions of master’s and doctoral students, this study uses a case narrative method to deeply describe the two CFL students’ academic emotional experiences’ changes and emotion regulation strategies of three feedback sessions in one semester, revealing the dynamics, richness, and complexity of academic emotions from multiple dimensions such as discrete emotions, activation, valence, and object focus. The results indicate that both positive and negative emotions have different promoting effects on the academic achievements of CFL students, helping them engage in writing learning and strive for academic improvement. Meanwhile, CFL students can also exert positive personality traits, regulate emotions, enhance resilience and well-being, and re-engage in the problem-solving process (Gross, 2015; Han and Xu, 2020; Oxford, 2016).

However, this study also has the following limitations. First, we only used writing ratings as measures of academic achievement. Future research could use more sophisticated measurement methods, such as lexical and syntactic complexity and accuracy, to gain a better understanding of the links between emotions and L2 writing achievement. Second, it is difficult to obtain and analyze CFL students’ emotional reactions to feedback through interviews and oral reports, and there may be problems such as inaccurate language expression, inaccurate emotion perception, and memory of L2 students.

Future studies could incorporate multiple measures to assess processing, such as combining traditional methods with tools such as eye-tracking, to gain a deeper understanding of feedback engagement. Additionally, the sample size and research environment of this study may limit the generalizability of the results. Despite these limitations, this study highlights the importance of exploring cultural and linguistic variability to enrich the application of positive psychology in language learning.

Future research could address the following aspects. For example, examining potential differences in learners’ values, expectations, and attitudes toward teacher feedback across different cultural backgrounds, analyzing how the linguistic characteristics and learning challenges of learners’ native languages and the target language (Chinese) influence emotional experiences. Such investigations would help ensure sample diversity and representativeness, particularly in terms of cultural background, thereby broadening the scope and applicability of findings in this field.

This study also sheds light on how academic writing teachers can provide feedback in a more acceptable and understandable way and how novice writers can use feedback to improve their academic writing skills. First, teachers should understand the complexity of learners’ emotions rather than assuming that feedback will inevitably trigger negative emotions that interfere with L2 writing learning. They should recognize that even negative feedback can trigger positive emotions and that learners can self-regulate negative emotions. Therefore, teachers need to explore how to use emotions in writing feedback, such as guiding students to express and reflect on emotions and increasing their awareness of the value of academic emotions.

Second, they could provide clear support for emotional issues triggered by academic writing feedback to help second-language learners regulate their emotions (Goetz et al., 2006). For example, they could help students establish a subjective control appraisal centered on self-agency, fostering a belief that they can improve their writing skills by actively engaging with feedback (Han and Xu, 2020). Teachers could also guide learners to make positive subjective value judgments about the importance of feedback, revisions, and writing tasks, helping students to regulate their emotions, and build psychological resilience, enhance their ability to benefit from feedback.

Finally, teachers also need to pay attention to the following points when giving feedback: First, they should provide guidance to students on how to review feedback so that they are not confused about the revision mode; second, they should pay attention to the effectiveness of feedback comments and strike a balance between positive and negative comments. This study found that negative feedback does not always evoke negative emotions while helping students to identify their own problems and make targeted improvements in future writing; positive feedback helps students to improve their self-agency in writing, recognize their own strengths, and develop them further. However, teachers generally focus on students’ language or structural

content problems, use negative comments, and ignore positive comments when giving feedback, which is not conducive to building students' confidence and enthusiasm for writing. If teachers strike a balance between correction, praise, and encouragement, would could effectively stimulate students' interest in writing and foster a positive cycle of academic writing and feedback.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Ethics statement

Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

RL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. PX: Methodology, Supervision, Validation, Writing – review & editing.

References

- Bruton, A. (2009). Improving accuracy is not the only reason for writing, and even if it were. *System* 37, 600–613. doi: 10.1016/j.system.2009.09.005
- Bruton, A. (2010). Another reply to Truscott on error correction: improved situated designs over statistics. *System* 38, 491–498. doi: 10.1016/j.system.2010.07.001
- Burić, I., Sorić, I., and Penezić, Z. (2016). Emotion regulation in academic domain: development and validation of the academic emotion regulation questionnaire (AERQ). *Personal. Individ. Differ.* 96, 138–147. doi: 10.1016/j.paid.2016.02.074
- Dewaele, J.-M., and Li, C. (2020). Emotions in second language acquisition: a critical review and research agenda. *Foreign Lang. World.* 1, 34–49.
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT J.* 63, 97–107. doi: 10.1093/elt/ccn023
- Ellis, R. (2010). "Cognitive, social, and psychological dimensions of corrective feedback" in *Sociocognitive perspectives on language use and language learning*. ed. R. Batstone (Oxford: Oxford University Press), 151–165.
- Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Q.* 31, 315–339. doi: 10.2307/3588049
- Foucault, M. (2012). Discipline and punish (revised edition). Trans. L. Beicheng and Yang Yuanyang. Beijing: SDX Joint Publishing Company 210–218.
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions. *Am. Psychol.* 56, 218–226. doi: 10.1037/0003-066X.56.3.218
- Gao, Z., and Yang, Y. (2023). The predictive effect of trait emotional intelligence on emotion regulation strategies: the mediating role of negative emotion intensity. *System* 112:102958. doi: 10.1016/j.system.2022.102958
- Geng, F., and Yu, S. L. (2024). Exploring doctoral students' emotions in feedback on academic writing: a critical incident perspective. *Stud. Contin. Educ.* 46, 1–19. doi: 10.1080/0158037X.2022.2109616
- Goetz, T., Frenzel, A., Pekrun, R., and Hall, N. (2006). "Emotional intelligence in the context of learning and achievement" in *Emotional intelligence: an international handbook*. eds. R. Schulze and R. D. Roberts (Cambridge, MA: Hogrefe and Huber), 233–253.
- Gross, J. J. (1998). The emerging field of emotion regulation: an integrative review. *Rev. Gen. Psychol.* 2, 271–299. doi: 10.1037/1089-2680.2.3.271
- Gross, J. (2015). Emotion regulation: current status and future prospects. *Psychol. Inq.* 1, 1–26.
- Gu, J. Y., Yang, D. L., and Lin, Q. P. (2022). A study of academic emotions and the corresponding regulation strategies in peer feedback in English writing from the perspectives of feedback providers and receivers. *Technol. Enhanc. Foreign Lang. Educ.* 4, 71–79.
- Han, Y., and Gao, X. S. (2024). Emotion regulation in L2 learning: theoretical underpinnings, current state of the research, and future directions. *Foreign Lang. Their Teach.* 1, 20–32. doi: 10.13458/j.cnki.flatt.004988
- Han, Y., and Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *J. Second. Lang. Writ.* 30, 31–44. doi: 10.1016/j.jslw.2015.08.002
- Han, Y., and Hyland, F. (2019). Academic emotions in written corrective feedback situations. *J. Engl. Acad. Purp.* 38, 1–13. doi: 10.1016/j.jeap.2018.12.003
- Han, Y., and Xu, Y. (2020). Emotional experiences and emotion regulation strategies in second language writing classrooms from the perspective of positive psychology: what can we learn from written corrective feedback situations? *Foreign Lang. World.* 196, 50–59.
- Han, Y., Xu, Y. T., Li, B. B., and Gao, X. S. (2024). Investigating postgraduate students' emotion regulation strategies involved in writing and publishing academic papers. *Modern Foreign Lang.* 1, 114–125. doi: 10.20071/j.cnki.xdwy.20231027.003
- Hyland, F. (1998). The impact of teacher written feedback on individual writers. *J. Second. Lang. Writ.* 7, 255–286. doi: 10.1016/S1060-3743(98)90017-0
- Hyland, K., and Hyland, F. (2006). Feedback on second language students' writing. *Lang. Teach.* 39, 83–101. doi: 10.1017/S0261444806003399
- Imai, Y. (2010). Emotions in SLA: new insights from collaborative learning for an EFL classroom. *Mod. Lang. J.* 94, 278–292. doi: 10.1111/j.1540-4781.2010.01021.x
- Jiang, Y., and Dewaele, J.-M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System* 82, 13–25. doi: 10.1016/j.system.2019.02.017

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1463488/full#supplementary-material>

- Jin, Y., and Zhang, L. J. (2018). The dimensions of foreign language classroom enjoyment and their effect on foreign language achievement. *Int. J. Biling. Educ. Biling.* 24, 948–962. doi: 10.1080/13670050.2018.1526253
- Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second language writing skills. *Mod. Lang. J.* 75, 305–313. doi: 10.2307/328724
- Kikuchi, K., and Lake, J. (2021). “Positive psychology, positive L2 self, and L2 motivation: a longitudinal investigation” in *Positive psychology in second and foreign language education*. eds. K. Budzińska and O. Majchrzak (Cham: Springer), 79–94.
- Kong, A., and Teng, M. F. (2023). The operating mechanisms of self-efficacy and peer feedback: an exploration of L2 young writers. *Appl. Linguist. Rev.* 14, 297–328. doi: 10.1515/applirev-2020-0019
- Lee, I. (2008a). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *J. Second. Lang. Writ.* 17, 144–164. doi: 10.1016/j.jslw.2007.12.001
- Lee, I. (2008b). Understanding teachers’ written feedback practices in Hong Kong secondary classrooms. *J. Second. Lang. Writ.* 17, 69–85. doi: 10.1016/j.jslw.2007.10.001
- Lei, J., and Hu, G. (2019). Doctoral candidates’ dual role as student and expert scholarly writer: an activity theory perspective. *Engl. Specif. Purp.* 54, 62–74. doi: 10.1016/j.esp.2018.12.003
- Li, C., Dewaele, J.-M., and Hu, Y. (2023). Foreign language learning boredom: conceptualization and measurement. *Appl. Linguist. Rev.* 14, 223–249. doi: 10.1515/applirev-2020-0124
- Lin, X. Y., and Lin, Q. Q. (2023). Excessive self-monitoring: the impact of assessment systems on top-notch innovative talents cultivation. *Glob. Educ.* 52, 14–32.
- Liu, C., and Yu, S. (2022). Exploring masters students’ emotions and emotion-regulation in supervisory feedback situations: a vignette-based study. *Assess. Eval. High. Educ.* 47, 1101–1115. doi: 10.1080/02602938.2021.2005770
- MacIntyre, P. D., Dewaele, J.-M., Macmillan, N., and Li, C. C. (2019). “The emotional underpinnings of Gardner’s attitudes and motivation test battery” in *Contemporary language motivation theory: 60 years since Gardner and Lambert (1959)*. eds. A. H. Al-Hoorie and P. D. MacIntyre (Bristol: Multilingual Matters), 57–79.
- Madhu, N. B., and Hu, G. (2021). Supervisory feedback across disciplines: does it meet students’ expectations? *Assess. Eval. High. Educ.* 46, 407–423. doi: 10.1080/02602938.2020.1780562
- Mahfoodh, O. (2017). “I feel disappointed”: EFL university students’ emotional responses towards teacher written feedback. *Assess. Writ.* 31, 53–72. doi: 10.1016/j.asw.2016.07.001
- McMartin-Miller, C. (2014). How much feedback is enough? Instructor practices and student attitudes toward error treatment in second language writing. *Assess. Writ.* 19, 24–35. doi: 10.1016/j.asw.2013.11.003
- Miles, M. B., Huberman, A. M., and Saldana, J. (2013). *Qualitative data analysis: a methods sourcebook*. 3rd Edn. Los Angeles, CA: Sage.
- Miles, M. B., Huberman, A. M., and Saldana, J. (2020). *Qualitative data analysis: a methods sourcebook*. 4th Edn. Los Angeles: Sage.
- Mirka, H., and Kirsí, P. (2019). Are learning skills associated with academic emotions elicited by Master’s thesis work? *J. Furth. High. Educ.* 43, 1299–1313. doi: 10.1080/0309877X.2018.1483012
- Oxford, R. L. (2016). “Toward a psychology of well-being for language learners: the “EMPATHICS” vision” in *Positive psychology in SLA*. eds. P. D. MacIntyre, T. Gregersen and S. Mercer (Bristol: Multilingual Matters), 10–87.
- Papi, M. (2010). The L2 motivational self- system, L2 anxiety, and motivated behavior: a structural equation modeling approach. *System* 38, 467–479. doi: 10.1016/j.system.2010.06.011
- Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications. *Educ. Psychol. Rev.* 18, 315–341. doi: 10.1007/s10648-006-9029-9
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., and Perry, R. P. (2011). Measuring emotions in students’ learning and performance: the achievement emotions questionnaire (AEQ). *Contemp. Educ. Psychol.* 36, 36–48. doi: 10.1016/j.cedpsych.2010.10.002
- Pekrun, R., Goetz, T., Titz, W., and Perry, R. P. (2002). Academic emotions in students’ self-regulated learning and achievement: a program of qualitative and quantitative research. *Educ. Psychol.* 37, 91–105. doi: 10.1207/S15326985EP3702_4
- Pekrun, R., and Linnenbrink-Garcia, L. (2012). “Academic emotions and student engagement” in *The handbook of research on student engagement*. eds. S. L. Christenson, A. L. Reschly and C. Wylie (New York: Springer), 259–282.
- Pitt, E., and Norton, L. (2017). “Now That’s the feedback I want!” students’ reactions to feedback on graded work and what they do with it. *Assess. Eval. High. Educ.* 42, 499–516. doi: 10.1080/02602938.2016.1142500
- Teng, M. F., and Ma, M. (2024). Assessing metacognition-based student feedback literacy for academic writing. *Assess. Writ.* 59:100811. doi: 10.1016/j.asw.2024.100811
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Lang. Learn.* 46, 327–369. doi: 10.1111/j.1467-1770.1996.tb01238.x
- Tugade, M. M., and Fredrickson, B. L. (2004). Resilient individuals use positive emotions to bounce back from negative emotional experiences. *J. Pers. Soc. Psychol.* 86, 320–333. doi: 10.1037/0022-3514.86.2.320
- Winstone, N., Boud, D., Dawson, P., and Heron, M. (2022). From feedback-as-information to feedback-as-process: a linguistic analysis of the feedback literature. *Assess. Eval. High. Educ.* 47, 213–230. doi: 10.1080/02602938.2021.1902467
- Yu, S., and Jiang, L. (2022). Doctoral students’ engagement with journal reviewers’ feedback on academic writing. *Stud. Contin. Educ.* 44, 87–104. doi: 10.1080/0158037X.2020.1781610
- Zheng, Y., and Yu, S. (2018). Student engagement with teacher written corrective feedback in EFL writing: a case study of Chinese lower-proficiency students. *Assess. Writ.* 37, 13–24. doi: 10.1016/j.asw.2018.03.001



OPEN ACCESS

EDITED BY

Xinghua Liu,
Shanghai Jiao Tong University, China

REVIEWED BY

Siti Drivoka Sulistyaningrum,
Jakarta State University, Indonesia
Arsen Nahum Pasaribu,
University of HKBP Nommensen, Indonesia
Raniya Alsehibany,
Riyadh College of Technology for Girls,
Saudi Arabia

*CORRESPONDENCE

Madalina Chitez
✉ madalina.chitez@e-uvt.ro

RECEIVED 21 October 2024

ACCEPTED 30 December 2024

PUBLISHED 20 January 2025

CITATION

Dudău DP, Chitez M and Sava FA (2025)
Elucidating the emotional persona in the
Romanian university students' academic
discourse: a corpus-based exploration.
Front. Psychol. 15:1514795.
doi: 10.3389/fpsyg.2024.1514795

COPYRIGHT

© 2025 Dudău, Chitez and Sava. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Elucidating the emotional persona in the Romanian university students' academic discourse: a corpus-based exploration

Diana Paula Dudău^{1,2}, Madalina Chitez^{3*} and Florin Alin Sava⁴

¹Department of Psychology and Psychotherapy, Faculty of Psychology, Titu Maiorescu University, Bucharest, Romania, ²Education Research Unit, National Center for Policy and Evaluation in Education, Bucharest, Romania, ³Department of Modern Languages and Literatures, Faculty of Letters, West University of Timișoara, Timișoara, Romania, ⁴Department of Psychology, Faculty of Sociology and Psychology, West University of Timișoara, Timișoara, Romania

Introduction: Despite growing global interest in the emotional dimensions of academic writing, Romanian academic discourse remains underexplored, particularly in multilingual contexts. This study addresses this gap by analyzing a bilingual corpus of texts written in Romanian (L1) and English (L2) across various disciplines and genres. It aims to uncover emotional dimensions conveyed through linguistic markers, exploring how language, culture, and academic context shape students' writing styles. Romania's historical and social emphasis on formality, hierarchy, and indirectness in communication serves as a backdrop for examining these dynamics.

Method: A corpus-based approach was adopted, utilizing the Linguistic Inquiry and Word Count 2015 (LIWC2015) tool to analyze linguistic and emotional markers. The bilingual ROGER corpus, containing texts from nine Romanian universities spanning multiple disciplines and genres, served as the dataset. Advanced data analysis techniques included supervised machine learning for language classification, network analysis to explore interactions among linguistic features, and cluster analysis to detect discipline- and genre-specific linguistic patterns.

Results: The findings reveal distinct emotional patterns between Romanian and English academic writing. Romanian texts exhibit a higher degree of formality and indirectness, while English texts reflect greater assertiveness and personal engagement. Additionally, the Romanian corpus demonstrates less linguistic cohesion and a broader range of writing styles. Genre- and discipline-specific trends also emerge, with English coursework and analytical writing, predominantly from social sciences, displaying more personal and emotional expression than research-focused texts. In contrast, the Romanian corpus, characterized by a third cluster, presents less clear-cut patterns: humanities texts span both emotionally expressive and neutral tones, while research and academic papers frequently exhibit an achievement-oriented or entrepreneurial style, though a significant subset also reflects a highly disengaged profile.

Discussion: By integrating machine learning, network analysis, and automatic language analysis, this study offers a novel perspective on how language, genre, and discipline-specific conventions shape emotional expression in academic writing. The results suggest that the Romanian students' emotional personas in academic writing are influenced by all these factors, potentially shaped by the

cultural norms of the second language, providing insights for teaching academic writing in multilingual settings.

KEYWORDS

academic writing, emotions, automatic language analysis, ROGER corpus, LIWC, cultural influences, multilingual higher education, cross-linguistic differences

1 Introduction

Academic writing is not only an educational skill that demonstrates the students' abilities to present, analyze, and communicate disciplinary content, but it also offers a window into their emotional and psychological states. This is particularly relevant for exploring whether features of academic discourse within a specific group reflect the emotional persona of that group. Pennebaker et al. (2014) demonstrated that subtle linguistic choices, such as function words in college essays, can reveal underlying cognitive and emotional processes, offering valuable insights into students' emotional engagement and academic success. In the case of Romanian university students, no prior research has been conducted to systematically identify the emotional prompts embedded in their academic discourse.

Investigating how Romanian students' writing reflects their emotions and attitudes is especially relevant, as it mirrors the societal shift from the communist era, where there was a tendency to conceal and repress thoughts, to the democratic period, where expressing opinions is both allowed and valued (Doroholschi et al., 2018). In addition, the introduction of additional writing cultures (Chitez and Kruse, 2012) into education, such as English-language norms, can contribute to significant changes in how students construct and express their identities. Exposure to different linguistic and rhetorical standards, particularly those that prioritize critical thinking and open discourse, encourages students to adopt more expressive and analytical approaches to academic writing. This cultural and linguistic shift not only broadens the students' communicative skills but also requires them to address the intricacies of expressing personal and emotional nuances within academic frameworks.

Despite the growing global interest in the emotional dimensions of academic writing, Romanian academic discourse remains underexplored. The legacy of collectivist educational practices from the communist era, which often emphasized conformity and formality, may have inhibited emotional self-expression. Moreover, adapting to international writing norms potentially without adequate pedagogical support might pose additional challenges for Romanian students, especially because all Romanian teachers and professors in activity since the fall of the communist regimen to present were educated in those times or were born right after the 1989 revolution. Thus, this dual tension – between preserving cultural identity and adopting global standards – adds complexity to elucidating emotional personas in students' writing and brings forward interesting questions.

Existing literature on the linguistic features of Romanian academic writing has primarily focused on phraseology (Chitez et al., 2021; Dincă et al., 2024; Muresan et al., 2022), argumentation (Tucan et al., 2020), and the development of computational resources such as the Romanian Academic Word List (Ro-AWL) (Bucur et al., 2022) and the Romanian Phrasal Academic Lexicon (ROPAL) (Chitez et al., 2021). These studies have offered valuable perspectives on both novice and expert academic writing, identifying key linguistic features that

shape Romanian academic discourse. Furthermore, contrastive analyses between Romanian and English academic writing datasets have revealed distinctive characteristics of the Romanian writing style, particularly in how argumentation is structured and phraseological units are employed (Manda and Chitez, 2022; Bercuci and Chitez, 2023). However, in previous Romania-specific studies, emotional and psychological elements are often overlooked despite their relevance to both academic performance and the understanding of larger societal values (Williams, 2017). This leaves a significant gap in understanding how Romanian students' linguistic choices reflect their emotional personas, and addressing this niche is crucial for developing targeted educational interventions in today's world shaped by globalization.

The relevance of the linguistics-driven psychological approach to academic writing is manifold. For instance, the scrutiny of the students' academic writing features, both linguistic and meta-linguistic, can reveal how they handle a disciplinary topic in terms of attitude: positively or negatively, assertively or hesitantly, confidently or with uncertainty (Hyland, 2005). The choice of words, sentence structure, and rhetorical strategies can indicate not only the level of subject mastery but also the emotional and psychological engagement of the writer (Hyland and Tse, 2007). For example, the use of modal verbs such as “might” or “could” may reflect hesitancy or a lack of certainty, while definitive language like “must” or “will” suggests assertiveness and confidence (Hyland, 2002). Additionally, variations in tone, whether formal, informal, or neutral, provide further clues to how students position themselves relative to the content, their audience, and the academic discourse community (Ivanič, 1998). These linguistic choices are often subconscious and can be influenced by a range of factors, including cultural norms, the perceived difficulty of the disciplinary field, and the expectations of the academic environment (Hinkel, 2001). Meta-linguistic features, such as hedging (e.g., “might,” “perhaps”), emphasis (e.g., “it is important to note”), boosters (e.g., “clearly,” “undoubtedly”), transition markers (e.g., “however,” “in addition”), frame markers (e.g., “first,” “finally”), and attitude markers (e.g., “unfortunately,” “interestingly”) play a crucial role in shaping academic writing by guiding the reader through the argument and indicating the writer's stance (Hyland, 1998). These elements help students navigate complex arguments, signal their engagement with the topic, and manage the relationship with their readers (Morita, 2004). Through careful analysis of these features, educators can not only correlate linguistic usage with the students' academic performance but also gain insights into their emotional and cognitive states, which are often interconnected with their writing decisions (Swales, 1990).

Adding to these complex aspects is the lack of validated tools for analyzing emotional and cognitive dimensions in multilingual settings that include Romanian academic writing. While there are many automatic language analysis tools capable of extracting emotional valence or contents from any text – for a review, see Eichstaedt et al. (2021) and Neuendorf (2017)—their application in the Romanian

language remains problematic because most of them were built for English and the translation and validation process of such instruments is not straightforward. Therefore, examining psychological markers of Romanian academic writing is in its infancy, which shows a pressing need for studies that bridge this research gap, especially concerning the emotional personas of Romanian students.

Building on this significant niche, in our study, we aim to address precisely this unexplored topic of elucidating the emotional persona in the Romanian university student's academic discourse. Using the Linguistic Inquiry and Word Count (LIWC) tool—the 2015 version developed by Pennebaker et al. (2015)—this research analyzes how Romanian students' writing in their native language (L1) and English (L2) reflects emotional and cognitive dimensions across various disciplines. This comparison between L1 and L2 texts enables us to investigate how language influences emotional and psychological expression in academic discourse, exploring whether students demonstrate different emotional personas when writing in their native language versus a second language. Furthermore, we assess how discipline-specific writing conventions influence the use of emotional and cognitive language, highlighting how academic fields shape students' linguistic and psychological expression. Through this, we aim to uncover both linguistic patterns and emotional markers that reflect the students' academic and personal identities.

Thus, the novelty of this research lies in its dual focus on emotional personas and multilingual academic writing. By applying LIWC2015, a validated tool for linguistic and psychological analysis, to a corpus of Romanian-English student texts, this study pioneers a rigorous approach to examining psychological markers in Romanian academic discourse. LIWC2015 was selected for its ability to extract a variety of psychological contents, including emotional, cognitive, and motivational dimensions, which could offer a comprehensive view of the emotional persona reflected in the Romanian student group's writing. Recently tested and validated for the Romanian language (Dudău and Sava, 2022), LIWC has proven its versatility in conducting detailed analyses of specific word categories (Kahn et al., 2007; Pennebaker et al., 2015). In addition, it features functionalities such as Keywords in Context (KWIC), which capture the nuanced aspects of academic writing (Tausczik and Pennebaker, 2010). By contextualizing specific words, LIWC captures the students' strategies for expressing analytical thinking, confidence, and emotional tone in their academic work. At the same time, LIWC's closed-vocabulary approach has shown immense research potential in comparative research across languages and disciplines (Kučera and Mehl, 2022). The Romanian LIWC also proved equivalence with other language versions, not only with the original English one (Dudău and Sava, 2021), allowing a future valid extension of our study to more intercultural comparisons of emotional personas emerging from language. Thus, in academic writing, where culture- and language-specific rhetorical choices are evident (Hinkel, 2002), LIWC provides powerful, user-friendly automated tools—a validated, multilingual dictionary and accompanying software—for systematically analyzing these differences. Its use in this study establishes a reliable foundation for investigating cross-linguistic and cultural variations in academic discourse, extending beyond Romanian and English, and further enhancing the rigor and relevance of our research.

In line with our objective to analyze the emotional persona in Romanian university students' academic discourse, this study aims to address the following key questions:

- 1 What are the key features of Romanian students' emotional personas as reflected in their academic discourse, particularly in terms of emotional expression (e.g., positive and negative emotions)?
- 2 How do these emotional personas differ between Romanian and English academic texts, and in what ways do these differences influence students' writing in each language?
- 3 What discipline-specific and genre-specific emotional identity traits can be identified in Romanian university students' academic writing?

2 Literature review

2.1 Emotional persona and academic writing: the path towards a socio-cognitive perspective

To reach an understanding of the complex dynamics between the writing process and the psychological factors correlated with it, research has undergone a dual model approach to writing: writing as a product (the final text) and writing as a process (the cognitive steps involved in creating that text). The beginnings of writing research focused on the perception of writing as a static outcome of the human mind. Early composition studies looked at text in terms of grammar, sentence accuracy, and formal structure without considering the cognitive processes associated with it. Writing was viewed in its developmental perspective as a progressive mastery of discourse types (Moffett, 1968) or as a craft that can be learned through applying linguistic rules and conventions (Murray, 1968). A writing research paradigm shift was signaled by Hairston (1982), who highlighted the need to view writing as a cognitive process since writing encompasses more than the mere application of linguistic rules. It involves complex mental stages such as idea generation, planning, drafting, and revising. Hairston (ibid.) argued that focusing solely on the final product neglects the recursive and dynamic nature of writing, where cognitive tasks interact with the writer's evolving text. This perspective paved the way for process-oriented approaches in writing pedagogy. Since the 1980s, this approach, as pioneered by Hayes and Flower (1980) and further developed by Bereiter and Scardamalia (2013), has conceptualized writing as a multifaceted interaction between cognitive functions, like working memory and executive control, and linguistic skills. As a result, the process approach has become the most comprehensive framework for understanding the complexity of writing (Alves and Haas, 2012), emphasizing how writers engage with their ideas and text throughout the writing process. Such view has greatly influenced writing pedagogy, advocating strategies that promote critical thinking and problem-solving skills in all writing activities, including academic writing.

Building on cognitive sciences, research has increasingly recognized that the analysis of emotional personas in academic discourse is deeply rooted in socio-cognitive perspectives on language and identity formation. These perspectives emphasize that writing is not only a reflection of individual cognitive processes but also a product of the social contexts in which it is produced. This aligns with Bereiter's (1980) view that the development of academic writing is closely linked to broader cognitive and developmental processes, such as social cognition and reflective thinking. Pohl (2007) further

highlights that key factors such as enculturation into academic norms, acquisition of disciplinary knowledge, and mastery of writing skills are shaped by both cognitive and social influences. In this view, writing is not merely a cognitive process but also a form of participatory sense-making that emerges through interaction with the social environment and artifacts (Vygotsky, 1978; Rogoff, 2003). The socio-cognitive perspective underscores that writing development is intertwined with socialization into academic norms and identity formation, thereby reflecting both cognitive functions and the socio-cultural contexts that shape academic discourse (Bereiter and Scardamalia, 2013; Pohl, 2007).

Such insights highlight that academic writing goes beyond technical skills, being a reflection of the writer's engagement with their social and intellectual environment. This perspective has expanded to include emotional personas as essential components of academic writing. Emotional personas are expressed through the writer's tone, style, and rhetorical choices, revealing their interaction with both the subject matter and the audience. Several studies have demonstrated the importance of examining both linguistic style and emotional expression to gain insights into the students' attitudes, confidence, and engagement. In a study on film reviews, Argaman (2010) demonstrated that emotions such as happiness or sadness are conveyed through linguistic choices, i.e., intensifiers, metaphors, and first-person pronouns, illustrating how these elements reflect the writer's emotional engagement with the content. In the case of academic writing, the emotional burden is heightened by the demands of the academic environment. In studies such as Negri et al. (2020), researchers have identified linguistic markers that carry a high emotional charge. Words such as "fear," "pain," and "despair" are indicative of heightened emotional arousal, signaling deep emotional responses to a topic. Cameron et al. (2009) examine how emotions such as self-doubt, anxiety, and fear are intricately connected to the challenges inherent in the research process. These emotions are further intensified by the critical nature of academia, as students "felt emotions like self-doubt, anxiety, and fear as shaped by the practices of critique" (Cameron et al., 2009, p. 274). This underscores how the critical framework of academia amplifies the emotional struggles faced by novice writers (ibid.).

In addition to the emotional labor involved in receiving and responding to feedback (Carless and Boud, 2018), academic writing involves several key emotional dimensions. Writers must balance asserting authority and expressing humility through linguistic strategies such as hedging (Hyland, 1996), the use of personal pronouns (I/we) (Hyland, 2002), and tone (Ivanič, 1998). Building on this, Liu (2013) study on the use of *Appraisal* resources in academic writing reveals that emotions such as satisfaction and personal engagement are expressed through authorial *Affect* values, which serve to project a strong personal voice in argumentative writing. These emotional cues, while subtle, play a crucial role in shaping the academic writer's identity and stance.

While international studies provide valuable insights, research on the socio-cognitive and emotional dimensions of academic writing in the Romanian context remains limited. As noted in the Introduction section, recent developments, such as the creation of resources like the Romanian Academic Word List (Ro-AWL) (Bucur et al., 2022) and the Romanian Phrasal Academic Lexicon (ROPAL) (Chitez et al., 2021), have advanced our understanding of linguistic features like phraseology and argumentation. However, the emotional aspects of academic

writing have been insufficiently explored, leaving a significant gap in understanding how Romanian students express emotions and attitudes through language, particularly when transitioning from their native language to English or another foreign language.

In December 2024, we conducted a search on the Web of Science Core Collection using keywords in the title that signaled language use (e.g., "language," "linguistic," "discourse," "writing," "text," "corpus," "phraseology") and keywords in the topic sections that reflected psychosocial variables (e.g., "persona," "emotion*," "cognit*," "attitudes," "motivation," "values"). The search was filtered to include studies analyzing both Romanian and English languages and focused on recent publications (i.e., the last 10 years). This yielded 42 documents, many irrelevant to our research focus, with only 19 addressing students. Among these, only three papers were tangentially (not directly) relevant to our study. For instance, one paper presented two corpora of business expressions in English and Romanian containing annotated metaphors suitable for cross-linguistic comparisons (Ferrari and Boca, 2017). In another study, Cojocaru (2021) analyzed 50 classroom compositions, revealing that several discourse markers (mainly textual connectors) differ between native Romanian speakers and students learning Romanian as a foreign language. Additionally, Senar et al. (2024) explored how the fluid intelligence of Romanian immigrant students shapes the relationship between L1 knowledge and L2 performance in Spanish and Catalan, showing some lexical, morphosyntactic, and orthographic particularities when speaking in Catalan versus Spanish.

The literature search also identified a few papers more linked to our research, even though they did not contain the word "students" or were not focused on academic writing. In line with our intention to capture changes in emotional expression between languages, Bromberek-Dyzman et al. (2021), testing two groups of bilinguals (Polish-English and Romanian-English), revealed cross-linguistic effects on emotional word recognition. On a different note, Popescu (2017) analyzed the metaphorical language in Romanian and British business press, detecting some notable differences, especially in attitudes towards work, whereas Ghivirigă and Baci (2015) showed that Romanian scientific texts demonstrate a preference for epistemic expressions through modal verbs, similarly to what previous literature on the English language indicated. Additionally, a few studies analyzed the discourse markers in Romanian and other languages to build a multilingual corpus (e.g., Silvano et al., 2022) or to investigate the linguistic borrowings in Romanian (e.g., Cojocaru, 2020), while others demonstrated efficient methods to establish a correspondence between English and Romanian metaphors or idioms despite socio-cultural footprints (Gogăță, 2023; Trantescu and Reiss, 2022). Finally, Boc (2020), in a theoretical paper, argued that language serves not only as a medium of communication but also as a determinant of national identity.

Despite these contributions, the lack of targeted research on how Romanian students' emotional personas adapt across languages remains evident. Understanding these adaptations requires a deeper investigation into the interconnections between cognition, language, emotions, and socio-cultural factors, underscoring the need for studies like ours.

2.2 The role of emotional persona in multilingual academic contexts

When students write in multiple languages, the emotional persona they project in their academic work may vary depending on the

cultural profiles shaped by their education and societal norms. These cultural imprints affect how they express emotions, assert authority, and engage with their audience, leading to different rhetorical choices and linguistic styles across languages. Kaplan (1966) and Cheng (1993) both explored how cultural thought patterns influence the structure of written discourse in the Chinese language, but they offered complementary insights into the topic. Kaplan proposed that Chinese writing often follows a circular or spiral thought pattern, characterized by indirectness and the gradual development of ideas. He suggested that Chinese students build their arguments by revisiting themes from different perspectives, which contrasts with the linear and thesis-driven structure typical in Western academic writing, such as writing in English. Cheng (1993), however, nuanced Kaplan's view by showing that while circularity and digressiveness are present, especially in introductions and conclusions, Chinese writing also incorporates linear elements. Cheng (ibid.) found that Chinese students use both deductive and inductive structures in body section types (i.e., initial / end and middle parts), resembling Western styles of argumentation in certain contexts. This blend of circular and linear approaches reflects the influence of both cultural traditions and modern academic conventions on Chinese students' writing. In his study of academic texts by L2 students from various linguistic backgrounds, Hinkel (2002) found notable differences in writing styles, influenced by students' first languages and cultural conventions. Chinese and Korean students often displayed more indirect argumentation, while Arabic speakers used elaborate, repetitive structures. Spanish-speaking students, instead, tended to write with more personal, subjective tones. These variations sometimes reflect the influence of different rhetorical traditions on L2 writing, indicating the challenges students face in adapting to English academic norms, particularly in terms of clarity and structure. Building on Kaplan's foundational ideas, Connor (1996) expands them by illustrating how English academic writing tends to be more linear and explicit in argumentation, while other cultures, such as Japanese or Arabic, might favor a more indirect or circular approach to presenting ideas.

Linguistic features can reveal specific aspects of writing cultures. Fløttum (2012) highlights notable differences in author visibility across academic writing in English, French, and Norwegian. English writers tend to use "I" more often, resulting in greater author presence and a more interactive style, where the writer frequently serves as a guide for the reader. In contrast, French academic writing employs the pronoun "on" (equivalent to "one" in English), which produces a more detached and abstract tone. Kruse et al. (2016) conducted an extensive analysis of academic writing in various European countries, offering valuable insights into the cultural and rhetorical factors shaping students' approaches. This broader exploration helps explain why students from different countries adopt diverse writing strategies, including the use of personal pronouns, stance, and hedging techniques.

From a contrastive rhetoric perspective, the Romanian writing style is a mixed type, sharing similarities and differences with other writing cultures. A study by Chitez and Kruse (2012) shows that Romanian academic writing is shaped by traditional educational practices that emphasize memorization and literature-based genres such as *comentariul literar* (literary commentary) and *analiza literară* (literary analysis). These genres foster formal, detailed argumentation, particularly in response to literary texts, which aligns with the country's teacher-centered system. However, educational policy shifts,

influenced by the Bologna process, have introduced internationally recognized genres like the opinion essay, posing challenges for students as they adapt to new writing norms without sufficient guidance. This evolution mirrors broader trends in Romanian writing culture, where traditional, national-specific genres are increasingly blending with global academic standards. The same has been demonstrated by Băniceru et al. (2012), highlighting the evolving influence of Anglo-Saxon writing norms on traditional Romanian academic writing. While Romanian writing was historically shaped by French academic models, focusing on descriptive elements and form, recent shifts reflect the adoption of more structured, concise, and reflective practices typical of Anglo-Saxon conventions. However, the transition is incomplete, as Romanian writing still prioritizes descriptive moves over critical analysis, suggesting a partial and mechanical integration of Western academic writing trends. In terms of linguistic features distinguishing Romanian natives' writing in Romanian versus English, several observations have been made. A corpus-based study by Bercuci and Chitez (2023) revealed that Romanian academic writing exhibits distinct linguistic traits that influence student writing, particularly when transitioning between Romanian and English. These include a preference for impersonal constructions and avoidance of first-person pronouns, which reflect a formal academic register. Romanian students tend to rely on descriptive and historicizing structures, frequently using phrases like "one of the most" and "at the same time," which are common in Romanian academic traditions. Additionally, the frequent use of prepositions (such as "de," "în," and "la" – "of," "in," and "to/at") and formulaic expressions indicates a focus on description and formality rather than argumentation. These features often carry over into English writing, where students struggle to adapt to the more concise, argumentative, and personal style expected in Anglo-Saxon academic norms. From an emotional persona perspective, such features are associated with formality, detachment, and indirect expression. This tendency may stem from cultural and educational traditions that prioritize respect for authority and objective reporting over direct, personal involvement in arguments. Consequently, Romanian students often show hesitancy in asserting personal opinions or taking ownership of their ideas, contrasting with the more assertive, individualistic style of English academic writing. This culturally rooted linguistic behavior can lead to challenges in achieving argumentative clarity and critical engagement when writing in English. However, no corpus-based analysis focusing exclusively on the emotional features of Romanian students' academic writing has been conducted.

In this context, the validation of the LIWC dictionary (Dudău and Sava, 2021, 2022) for use in academic research is a valuable tool, as it allows for detailed linguistic comparisons between Romanian and English texts. By enabling researchers to systematically analyze language use across these two languages, this validated dictionary supports the exploration of key linguistic features such as emotional tone, formality, and complexity.

2.3 Discipline-specific and genre-specific emotional identity traits

Academic writing is not only influenced by socio-cognitive factors and language- or culture-specific rhetorical traditions but also by the

disciplinary and genre conventions that shape how knowledge is communicated within a field. Numerous studies have shown that the disciplinary epistemologies, communication patterns, and discursive practices differ from discipline to discipline (for instance, Langer and Applebee, 1987; Bazerman and Paradis, 1991; Monroe, 2002; Poe et al., 2010; Thaïs and Myers Zawacki, 2006). The variation in disciplinary conventions is due to differences in knowledge production, rhetorical goals, and audience expectations (Hyland, 2004). In hard sciences, writing is objective, concise, and data-driven, focusing on clarity and empirical evidence (Varttala, 2001), while humanities and social sciences make use of figurative language and demonstrate deeper engagement with sources to create emotional resonance and nuanced meaning (Machin and Mayr, 2012). Varttala (2001) also found that the use of hedging, or cautious language, varies across different disciplines, including economics, medicine, and technology. Citation practices also differ, with scientific fields favoring concise references to current research (Hyland, 1999), and humanities offering extended commentary on sources (Swales, 1990).

However, in point of the emotional approach to writing, numerous recent sentiment analysis studies have identified a generalizing trend called linguistic positivity bias, first discussed in research by Vinkers et al. (2015), which explored the use of positive and negative words in scientific PubMed abstracts between 1974 and 2014, showing that positive language increased more rapidly than negative language. In line with this, Xiao et al. (2023), who examined the evolution of sentiment in academic writing in China across the humanities and social sciences over time, found a noticeable shift towards more positive sentiment in recent decades. A study by Chen (2024) confirmed the distinct tone of medical writing, particularly in how it conveys emotions such as trust, hope, and surprise when addressing groundbreaking discoveries or unexpected findings. These emotions are subtly embedded through careful word choice – positive framing is used to highlight successful outcomes, while more cautious or measured language is employed when discussing study limitations, creating a balance between excitement and professionalism in medical discourse.

Specific emotion-signaling linguistic strategies are also genre-specific. The work of Swales (1990, 2004) pioneered the analysis of research genres and made the language of research accessible to scrutinized study. Swales work was a milestone in the study of research genres and in introducing methods from applied linguistics to the study of English as a research language. His corpus approach has been picked up by other researchers like Hyland (2000, 2005, 2008, 2009, 2012), who engaged in systematic corpus studies on such issues as metadiscourse, citation signals, praise and criticism, power and authority, use of “I”/“we.” This type of research is complementary to the concept of emotional persona in academic writing, as it examines how linguistic choices, such as personal pronouns, tone, and metadiscourse, reveal the writer’s emotional engagement, confidence, or detachment in scholarly discourse.

Previous research has shown that the linguistic cues present in academic writing provide valuable insights into how emotions such as positivity, enthusiasm, uncertainty, or confidence are conveyed within academic discourse. These cues offer a deeper understanding of the writer’s academic experience. Ultimately, academic writing is shaped by the interaction of cognitive processes, personal emotions, and the social and cultural norms of the academic community. Analyzing large linguistic datasets allows researchers to identify patterns of

emotional expression and the rhetorical strategies employed by specific groups of writers.

3 Method

3.1 Corpus

For this study, the source of student writing was ROGER, a bilingual corpus of academic texts collected in 2018–2021 within nine Romanian universities (Chitez et al., 2021). As depicted in the ROGER platform (Strilețchi et al., 2022), the corpus contains 1,139 texts in English and 911 in Romanian, spanning various genres and being written by students at the Bachelor’s, Master’s or PhD degree levels across eight disciplines. The ROGER corpus was selected because it captures real-world academic writing produced by Romanian students across diverse disciplines, academic levels, and genres, thereby enhancing the generalizability of findings to a broader context of Romanian academic discourse. Initiated in 2017, it was the first bilingual Romanian-English learner corpus of this nature (Oravițan et al., 2022). To the best of our knowledge, the ROGER corpus offers a unique resource for studying academic writing within the Romanian context.

To reduce the class imbalances in genre and discipline, which could bias the results of the data analysis due to the overrepresentation of certain categories, we preprocessed these two categorical variables. In this vein, genres were grouped into two main categories: (1) *coursework and analytical writing*, encompassing essays, literary analyses, reviews, summaries, reading notes, assignments, tutorials, paragraphs, portfolios, CVs, interviews, and letters; (2) *research and academic papers*, comprising research papers, reports, Bachelor’s theses, Master’s theses, projects, and project documentations. This distinction was meant to separate reflective or summarizing tasks that allowed for more personalized language from formal, more standardized academic writing, which is typically used in research papers or other specialized materials. Similarly, the discipline variable was reduced from eight to three categories by combining the texts from computer science, engineering, and mathematics into *STEM*, those from political science, social science, economics, and law into *social sciences* while keeping *humanities* as a standalone class.

3.2 Automatic language analysis

3.2.1 Tools and linguistic variables

To extract linguistic content and style from the ROGER texts, we used Linguistic Inquiry and Word Count 2015 (LIWC2015). The original English version developed by Pennebaker et al. (2015) was applied to the English texts, while the Romanian adaptation (Ro-LIWC2015; Dudău and Sava, 2021, 2022) was used for the Romanian texts. LIWC2015 is a closed-vocabulary text analysis tool consisting of a piece of software capable of determining the percentage of words in the input texts based on over 90 grammatical and psychological categories defined in a so-called *dictionary*, a list of labeled words, word stems, and emoticons established through rigorous research. The English LIWC2015 dictionary (Pennebaker et al., 2015) contains 6,549 entries, while the Romanian one includes 47,825. This difference

in length is due to the particularities of Romanian in terms of morphology, semantics, and diacritics compared to English. Nevertheless, validation studies for RO-LIWC2015 (Dudău and Sava, 2021, 2022) have demonstrated that both dictionaries produce comparable results, indicating compatibility across languages.

From the multitude of LIWC2015 variables, we selected the following subset, which we considered most relevant to the goals of the current study: (1) first- and second-person pronouns (*i*, *we*, and *you* categories), as they indicate where the communication is directed—whether self-centered, toward a group with which the author identifies, or addressed to another person(s); (2) several parts of speech—*articles*, *prepositions*, *adverbs*, *conjunctions*, and *adjectives*—that suggest the degree of elaboration or complexity in the texts' structure; (3) *verbs*, which show the extent to which the texts are action-oriented; (4) *positive* and *negative emotions*, as indicators of affective valence; (5) *family* and *friend* categories, illustrating a focus on close social relationships; (6) cognitive processes—*insight*, *causation*, *discrepancy*, *tentative*, *certainty*, and *difference*—that outline the depth of thinking; (7) motivational drives—*affiliation*, *achievement*, *power*, *reward*, and *risk*—that reflect key forces guiding behaviors or perspectives; (8) time orientation, showing whether the texts focus on the *past*, *present*, or *future*; (9) personal concerns—*work*, *leisure*, *home*, *money*, *religion*, and *death*—that reveal the presence of topics related to major life domains.

3.2.2 Text selection and final dataset

Since there is no universally established minimum word count for valid text analysis with LIWC2015, we initially adopted the criteria used by Boyd and Schwartz (2021) to test the psychometrics of the LIWC-22 dictionary. Accordingly, we selected texts from the ROGER corpus that contained at least 100 words and had at least 65% of the words covered by the LIWC2015 dictionary (in English or Romanian, depending on the language of the text). However, applying these criteria resulted in the exclusion of about 27% of the Romanian texts, many of which were written in highly specialized language.

Excluding such a large portion of texts could have disproportionately affected the representation of certain genres or disciplines, potentially undermining the validity of our dataset. Therefore, to retain more valuable data without compromising the quality of the analysis, we adjusted the coverage threshold to 60% while maintaining the 100-word minimum. This adjustment allowed us to include 88.6% of the Romanian ROGER corpus and 98.3% of the English ROGER corpus, ensuring that a sufficient portion of each text's linguistic data was analyzed for meaningful results. The final dataset for our study, following this selection, is presented in Table 1.

As Table 1 indicates, there was a notable difference in LIWC2015 dictionary coverage between the English (83.14%) and Romanian (69.48%) texts. This difference might be attributed to a combination of factors, but the most prominent one could be that English was a foreign language for most of the students who wrote the ROGER texts (approximately 94% of the selected texts were written by Romanian students). Therefore, they may have used simpler, more general vocabulary, which is better represented in the LIWC2015 dictionary. In contrast, the Romanian texts, written in the students' native language, may contain more specialized or nuanced academic terminology, which is likely less covered by the Romanian LIWC2015 dictionary.

TABLE 1 Composition and linguistic characteristics of the final ROGER subset used in this study.

Composition	English corpus	Romanian corpus
Number of texts	1,120	807
Discipline		
Stem	368	37
Social sciences	474	214
Humanities	278	556
Genre		
Coursework and analytical writing	770	628
Research and academic papers	350	179
LIWC2015 tokenizer statistics		
Word count— <i>m(sd)</i>	1,782.51 (3,870)	1,374.32 (2,587.42)
Words per sentence— <i>m(sd)</i>	25.19 (10.18)	25 (10.70)
Dictionary coverage— <i>m(sd)</i>	83.14% (6.43)	69.48% (4.95)

Moreover, as shown in Table 1, the ratio between the English and Romanian texts is roughly 1.4 to 1, a moderate imbalance that would not necessarily require special attention during data analysis. In contrast, there were significant imbalances by genre and discipline, and we analyzed the linguistic markers associated with these two variables using a different approach, as explained in Section 3.3.

3.3 Data analysis strategy

To uncover the linguistic style and psychological contents in student academic writing and to reach more nuanced interpretations, we adopted a three-pronged approach, with each dimension complementing the others: (1) distinguishing between Romanian and English in student writing; (2) exploring the interactions between linguistic features; and (3) uncovering linguistic patterns. Throughout these analyses, we used different machine learning and statistical methods to provide multiple perspectives and deepen our understanding of student academic writing. Additionally, where appropriate, we applied cross-validation to manage the bias-variance trade-off, thereby improving the reliability and generalizability of our interpretations. Given that ROGER is a bilingual corpus, we performed within-language standardization before any data analysis. Specifically, for each LIWC2015 variable, we computed *z*-scores based on the mean and standard deviation of each language subsample, as suggested by previous research on multilingual data (Dudău and Sava, 2021; Meier et al., 2018). The following paragraphs provide detailed explanations of these technical aspects.

For the first objective—testing whether there are linguistic differences between Romanian and English languages in student writing—we applied two machine learning algorithms: *logistic regression* and *random forest*. Both addressed the classification problem of detecting language (English versus Romanian) based on the linguistic style and psychological contents assessed with LIWC2015. We started with logistic regression because it is a widely used and interpretable classification method that effectively detects linear relationships between the predictor variables and a binary

outcome. Then, we built a random forest model, as this algorithm, by growing multiple de-correlated decision trees and averaging their predictions (Breiman, 2001; Hastie et al., 2009), is able to capture potential non-linear relationships between input and output and complex interactions between the linguistic features.

To ensure the robustness of the classification models, we implemented cross-validation for two purposes: to test the models on unseen data and to tune the random forest model. Specifically, we employed a stratified train-test split, selecting 75% of the data for training and 25% for testing while preserving the proportion of Romanian and English texts in both subsets. After the split, we performed within-language standardization on the LIWC2015 variables in the training subset. The z-scores were calculated separately for each language subset, using the mean and standard deviation of the respective subset. The same transformation was then applied to the test subset (i.e., the z-scores for the test subset were computed based on the means and standard deviations on the training subset to prevent data leakage and keep the test data exclusively for assessing the model performance).

For tuning the random forest model, we used 10-fold cross-validation and two accuracy metrics – area under the ROC curve (AUC) and F1-score – to evaluate the performance of different hyperparameter combinations. This cross-validation method involved dividing the training subset into ten equal folds, training the model on nine folds, and validating it on the remaining fold. The process was repeated ten times, with each fold used once as the validation set. We focused on four hyperparameters: the number of trees in the forest, the number of LIWC2015 features randomly selected at each split, the minimum number of texts in a leaf, and the maximum number of leaves. For each hyperparameter, we defined a search space: the number of trees ranged from 100 to 1,000, the number of predictors from 1 to 33, the node size from 1 to 20, and the maximum nodes from 10 to 100. A random search method, iterating over 500 combinations of these hyperparameters, was employed to identify the optimal combination based on the highest mean accuracy in the 10-fold cross-validation process.

After building the machine learning models on the training subset, we assessed the classification accuracy on the test subset. In this regard, multiple parameters were computed—accuracy, sensitivity (true positive rate), specificity (true negative rate), F1-score, and AUC. The higher these values, the better the classification accuracy. For AUC, clear benchmarks exist to aid in interpretation: AUC values between 0.50 and 0.70 are generally considered to show low accuracy, values between 0.70 and 0.90 indicate moderate accuracy, and values above 0.90 suggest high accuracy (Akobeng, 2007). In our study, achieving at least moderate classification accuracy suggested the presence of notable linguistic differences between the English and Romanian corpora, with higher accuracy indicating more pronounced distinctions.

To address our second data analysis objective—exploring the interactions between linguistic features—we conducted a network analysis using the 33 LIWC2015 variables as nodes. Before implementing this approach, we performed within-language standardization. Network analysis is particularly valuable when elements of interest can be viewed as components of a system where each is connected to others (Borsboom et al., 2021). Considering that natural language consists of words linked through semantic, morphological, and syntactic rules, which might resemble a system, network analysis can provide a novel perspective on student writing through the lens of linguistic features.

Specifically, to model the relationships between LIWC2015 variables and identify key linguistic interactions, we estimated a Gaussian graphical model using graphical LASSO regularization combined with the extended Bayesian information criterion (EBIC) for edge selection, following guidelines from Epskamp et al. (2018). This approach produces a parsimonious network where edges represent partial correlations between variables, accounting for all other variables in the analysis. The choice of LASSO regularization with EBIC was made to ensure that our network focused on the most prominent linguistic connections, balancing interpretability with accuracy. While this method has high specificity, meaning it effectively removes non-existent edges, it may be less sensitive in detecting true edges (Epskamp and Fried, 2018). Given the exploratory nature of this approach to academic writing, we prioritized interpretability, even if it meant potentially excluding some true edges. After estimating the network structure, we computed four centrality measures—betweenness, closeness, strength, and expected influence—to identify the most influential linguistic features in the network, providing insights into how these features interact and shape student writing. Finally, we assessed the stability of the network using bootstrap methods.

To meet the third and final data analysis objective – uncovering linguistic patterns across genres and disciplines—we applied k-means clustering, an unsupervised learning algorithm. The same set of 33 LIWC2015 categories was used as input variables for this analysis, preceded by within-language standardization to ensure comparability between the English and Romanian corpora. Specifically, k-means clustering allowed us to explore whether distinct types of texts emerged based on their linguistic features. To determine the optimal number of clusters, we used the majority rule method, testing solutions with 2 to 15 clusters. This method evaluates several cluster validity measures and recommends the number of clusters supported by the majority of these indices (Lesmeister, 2015). The clusters were built based on 1,000 random starting sets. Ultimately, to uncover potential linguistic differences across genres and disciplines, we applied the Chi-squared test to examine whether the cluster distribution was significantly associated with the texts' genres and disciplines.

All analyses described in this section were performed using R and RStudio. Data manipulation and visualization were carried out using the *tidyverse* package (Wickham et al., 2019). The stratified train-test split was implemented with the *caTools* package (Tuszynski, 2021). Logistic regression was performed using the *glm* function from R's base package, while the random forest model was trained and evaluated within the *mlr* framework (Bischl et al., 2016). Network estimation, visualization, and description were facilitated by the *qgraph* package (Epskamp et al., 2012), while network stability was assessed using the *bootnet* package (Epskamp et al., 2018). For k-means clustering, we used R's built-in *kmeans* function from the *stats* package, in conjunction with the *NbClust* package (Charrad et al., 2014) for determining the optimal number of clusters.

4 Results

4.1 Distinguishing between Romanian and English languages in student writing

Altogether, the two classification models—logistic regression and random forest – used to differentiate between academic texts written

in English and Romanian, based on the 33 LIWC2015 features, revealed complex distinctions between the two corpora.

Specifically, the performance of the logistic regression model on the test data was poor. Predicted probabilities for language classification were nearly constant and consistently below 0.50, leading to the misclassification of all texts as Romanian. This resulted in a low AUC of 0.45, which falls well below the commonly accepted threshold of 0.70 for acceptable classification accuracy. On the training subset, although the model's intercept was significant ($\beta = -0.33$, $SE = 0.05$, $z = -6.15$, $p < 0.001$), all predictor variables had p -values of 1, indicating no significant contribution to the model. The null deviance (1964.8) and residual deviance (also 1964.8) suggest that the inclusion of the LIWC2015 predictors did not significantly improve the model over a null model. Multicollinearity was not a major concern, as most variance inflation factor (VIF) values were below 5, except for two variables: *verbs* (VIF = 6) and *focus on the present* (VIF = 5.24). However, these values are still not alarming, as VIF values below ten are generally considered acceptable (Bowerman et al., 2015; Field, 2018), and some scholars suggest that even higher values may not justify the exclusion or preprocessing of some variables (O'Brien, 2007). These findings suggest that any linguistic differences between the Romanian and English texts, if present, were likely subtle or involved non-linear relationships, which logistic regression cannot capture effectively.

In line with this observation, the random forest model, which is better suited for detecting complex and non-linear patterns, displayed excellent classification accuracy on the test subset, with performance parameters close to 1, as depicted in Table 2. The tuning process yielded the optimal parameters of 716 trees, 2 predictors randomly selected at each split, a minimum node size of 7, and a maximum of 94 terminal nodes.

The top 10 most relevant linguistic features distinguishing Romanian from English student writings were the word frequencies for *death*, *home*, *family*, *religion*, *I*, *friend*, *we*, *you*, *money*, and *leisure*. Table 3 presents the entire feature hierarchy based on the *mean decrease Gini* value, which indicates how important each feature was in reducing the impurity of the trees in the random forest model, with higher values reflecting greater importance.

Notably, personal concerns (except *work*) and personal pronouns dominated the top 10 list, alongside the social categories *family* and *friend*. Although direct comparisons of word percentages between languages are not the primary focus of random forest models, we observed that, on average, Romanian texts contained far fewer first-person pronouns, both singular and plural (see Table 3). This might suggest less personal engagement in Romanian writing compared to English. Therefore, a simplified inference might be that Romanian texts tend to exhibit a more formal style than the ones written in English.

Given the nearly perfect classification accuracy of the random forest model, we decided to keep the Romanian and English corpora

separate for the remaining analyses, allowing us to explore language-specific linguistic patterns in greater depth.

4.2 Exploring the interactions between linguistic features in student writing

To further investigate the relationships between linguistic features in student writing, we conducted separate network analyses for the Romanian and English corpora, utilizing the 33 LIWC2015 features as nodes in the network, as discussed in section 3.3 *Data analysis strategy*. The goal was to identify how linguistic features interact and contribute to the overall structure of student writing. Figure 1 provides the visual representation of the two parsimonious networks of partial correlations between linguistic variables.

As observed, the network for the English texts displayed higher interconnectedness than the Romanian network, which was confirmed by the degree centrality metric. Degree centrality reflects the number of connections (or direct relationships) each node has. Specifically, in the English corpus, the number of connections per node ranged from 12 to 23, with the *power* category showing the highest number of connections, while *achievement*, *certainty*, and *money* had the fewest. In contrast, the Romanian corpus network exhibited fewer connections per node, ranging from 3 to 14. The second-person pronoun (*you*) category had the fewest connections, whereas *discrepancy*, *focus on the present*, *money*, *positive emotions*, and *work* had the highest number of connections.

To gain further insights into the linguistic interactions within each corpus, we computed four additional centrality metrics: strength (the sum of the absolute edge weights connected to a node), closeness (the inverse of the sum of all distances from a node to all other nodes, with higher values indicating closer proximity to the entire network), betweenness (how often a node lies on the shortest path between any two other nodes), and expected influence (a measure similar to strength but taking into account the direction of connections, with negative correlations reducing the influence of a node) as defined by Deserno et al. (2022). Figure 2 presents the results for the English corpus, and Figure 3 shows the corresponding results for the Romanian corpus.

For the English network, the strength metric revealed that *verbs* played a pivotal role in shaping the structure of English texts, strongly connecting to other linguistic categories. Other linguistic variables that could directly affect or be affected by many writing characteristics were *work*, *focus on the present*, *home*, and *first-person singular pronouns*. Similarly, the expected influence metric showed that *verbs*, *affiliation* drive, *achievement* drive, *positive emotions*, and *reward* drive emerged as the most influential variables, shaping the overall structure of the linguistic network. According to the closeness metric, the use of *first-person singular pronouns*, words referring to *home*, *discrepancy*, *work*, and *risk*, as

TABLE 2 Performance metrics for logistic regression and random forest models in detecting the language of texts.

Model	Accuracy	Sensitivity	Specificity	F1-score	AUC
Logistic regression	42%	1	0	0.59	0.45
Random forest	99%	1	0.99	0.99	0.99

The results were obtained using the test subset ($n = 482$), with "Romanian" as the positive class.

TABLE 3 Importance of LIWC2015 features in the random forest model for identifying the language of texts, with feature-level means and standard deviations (prior to standardization).

Random forest model			Descriptive statistics— <i>m(sd)</i>	
Hierarchy of LIWC2015 features		Mean decrease gini	Romanian texts	English texts
1.	Death	86.47	0.17 (0.35)	0.20 (0.43)
2.	Home	80.72	0.12 (0.23)	0.51 (1.33)
3.	Family	70.31	0.13 (0.33)	0.22 (0.46)
4.	Religion	64.74	0.24 (0.59)	0.24 (0.57)
5.	I	58.33	0.18 (0.43)	1.44 (2.27)
6.	Friend	54.53	0.15 (0.21)	0.20 (0.38)
7.	We	37.06	0.33 (0.42)	1.11 (1.60)
8.	You	35.21	0.26 (0.32)	0.63 (1.21)
9.	Money	19.52	0.76 (1.21)	1.12 (1.55)
10.	Leisure	18.55	1.02 (1.16)	0.99 (1.25)
11.	Focus on the future	12.38	0.46 (0.45)	1.03 (0.96)
12.	Risk	9.20	0.92 (0.80)	0.76 (0.81)
13.	Discrepancy	9.12	1.49 (0.93)	1.79 (1.40)
14.	Work	8.17	4.32 (2.46)	4.51 (3.46)
15.	Negative emotions	7.00	2.12 (1.47)	1.49 (1.18)
16.	Focus on the past	6.84	5.96 (2.18)	2.42 (1.60)
17.	Positive emotions	6.68	3.49 (1.62)	3.07 (1.56)
18.	Reward	6.06	1.05 (0.84)	1.19 (0.82)
19.	Achievement	5.60	2.87 (1.69)	2.00 (1.22)
20.	Tentative	5.45	3.07 (1.33)	2.49 (1.23)
21.	Articles	5.38	4.61 (1.35)	9.41 (2.53)
22.	Affiliation	5.38	1.58 (1.10)	2.33 (1.80)
23.	Verb	4.96	12.40 (2.28)	1.19 (0.82)
24.	Insight	4.91	4.04 (1.56)	2.75 (1.21)
25.	Prepositions	4.89	13.80 (1.77)	14.60 (1.79)
26.	Conjunctions	4.84	4.81 (1.52)	6.33 (1.40)
27.	Causation	4.83	3.75 (1.46)	2.58 (1.20)
28.	Focus on the present	4.62	6.68 (1.86)	8.66 (2.63)
29.	Difference	4.59	2.98 (1.22)	2.78 (1.21)
30.	Power	4.54	3.67 (1.81)	2.98 (1.57)
31.	Adverbs	4.52	6.30 (1.75)	3.50 (1.34)
32.	Certainty	4.26	2.00 (0.84)	1.46 (0.78)
33.	Adjectives	4.15	7.03 (1.77)	4.78 (1.40)

The results were obtained using the training subset ($n = 1,445$), with “Romanian” as the positive class.

well as the frequency of *verbs*, ensured information flow within the network, having a high probability of being easily affected when another linguistic feature changed in the network. Additionally, betweenness values were highest for *verbs*, *home*, *first-person singular pronouns*, *work*, and *discrepancy*, indicating that these linguistic features served as key connectors, bridging otherwise

disparate elements in writing and facilitating transitions between different ideas or topics.

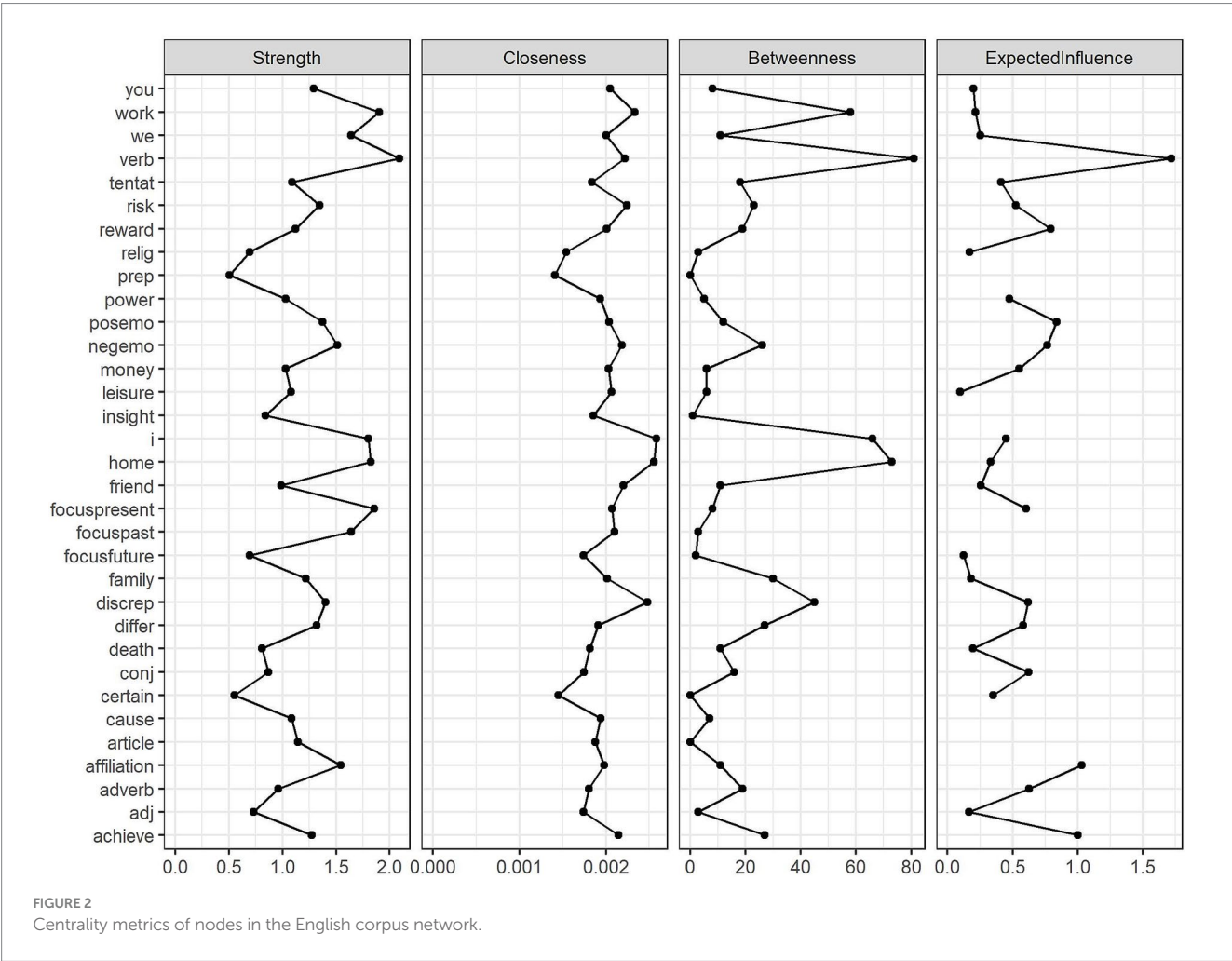
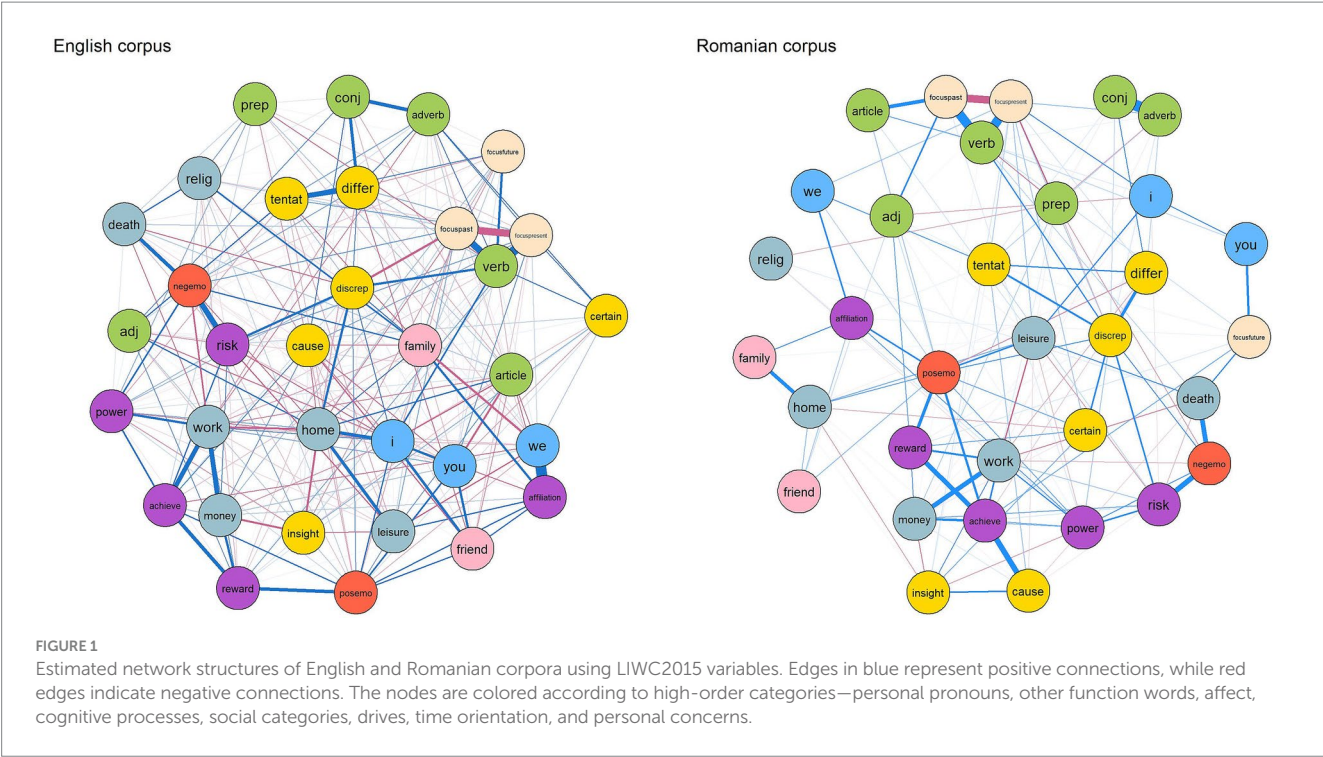
As far as the Romanian network was concerned, *focus on the present*, *verbs*, *focus on the past*, *achievement*, and *positive emotions* had the highest strength, indicating that action- and present-oriented language was about as central in Romanian writing as in the English texts. Likewise, *verbs* led in expected influence, followed by *achievement*, *positive emotions*, *discrepancy*, and *reward*. The least peripheral linguistic categories in the Romanian corpus network were *discrepancy*, *positive emotions*, *reward*, *achievement*, and *leisure*, meaning they acted as bridges between various linguistic features. The high closeness of *discrepancy*, in particular, may suggest that Romanian student writing, like English writing, contains nuanced or contrasting language to transition between ideas. The other variables high in closeness might indicate that the flow of information in Romanian student writing might be sustained mostly by addressing positive topics. In terms of betweenness, the linguistic categories measuring the focus on *positive emotions*, *discrepancy*, *present time*, *achievement*, and *risk* were prominent, suggesting that removing these linguistic markers would significantly disrupt the connections between other linguistic features, further highlighting their bridging role in Romanian academic writing.

After the network estimation step, we used two bootstrap methods to assess the stability of the estimated networks. Specifically, we computed 95% confidence intervals for the edge weights (see Figure 4). Overall, both networks demonstrated multiple strong and reliable connections with narrow confidence intervals, though some edges were weak or potentially unstable. The case-dropping bootstrap method, which evaluates how the network structure changes when portions of the data are removed, was applied to assess the stability of the strength centrality across the networks. This method showed that even when up to 50–70% of the data was excluded, the strength centrality measures remained highly correlated with the full-sample estimates (see Figure 5).

4.3 Uncovering the linguistic patterns in student writing across genres and discipline

At the final stage of our data analysis, we applied the k-means clustering algorithm followed by the Chi-square test to examine whether distinct linguistic markers, based on the 33 LIWC2015 categories, organized the texts into meaningful groups and whether these groups varied by genre and academic discipline. The analysis was conducted separately for the English and Romanian corpora, using 1,000 random starting sets. According to the majority rule method, the English corpus was best represented by two clusters, whereas the Romanian corpus was represented by three.

Table 4 outlines the characteristics of these clusters through the mean z-scores for each LIWC2015 category. Although the within-language standardization procedure made the variables lose their original meaning (word percentages), we decided to use it before the k-means clustering, too, because Romanian and English might inherently have different linguistic distributions for certain LIWC2015 categories due to grammatical particularities and our interest lay in the linguistic patterns shaped by psychological or cultural factors in student writing.



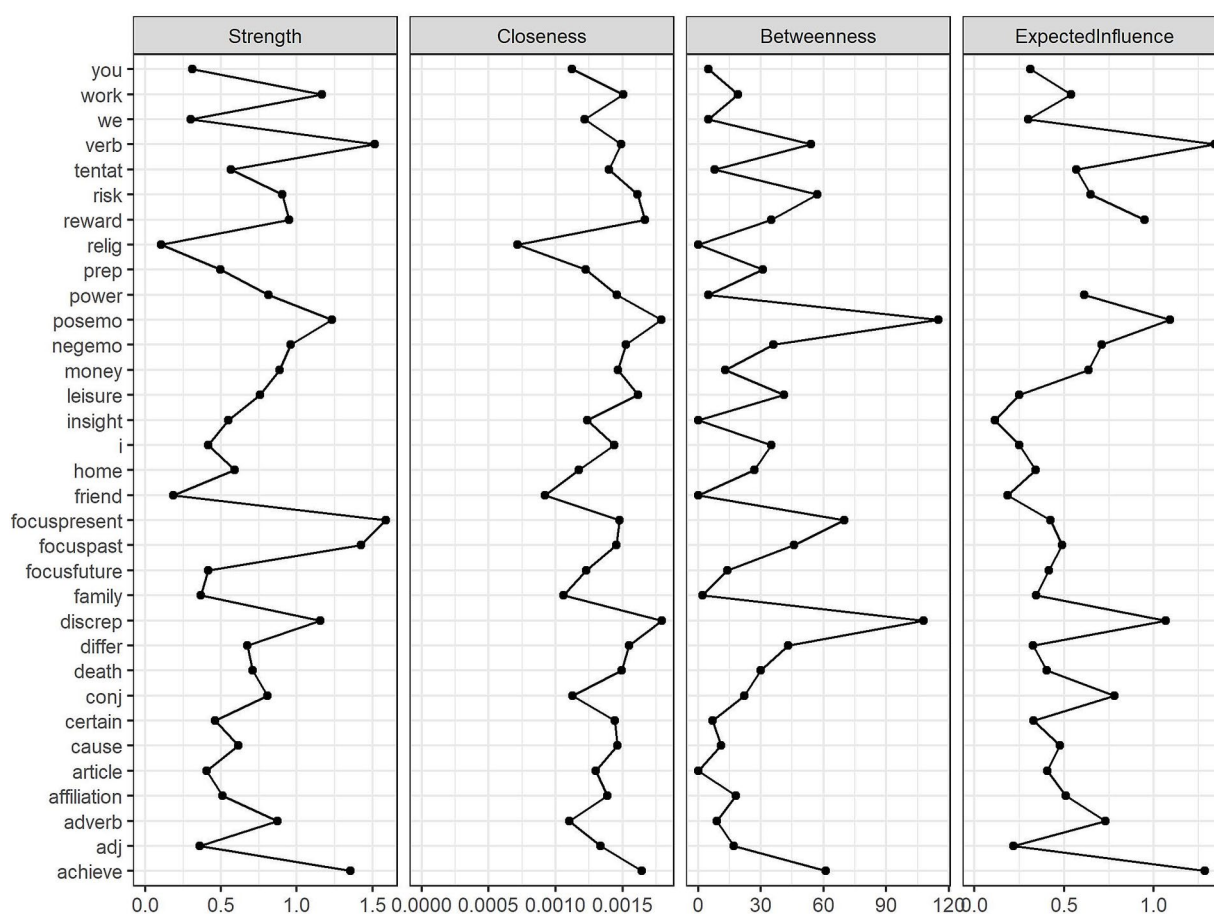


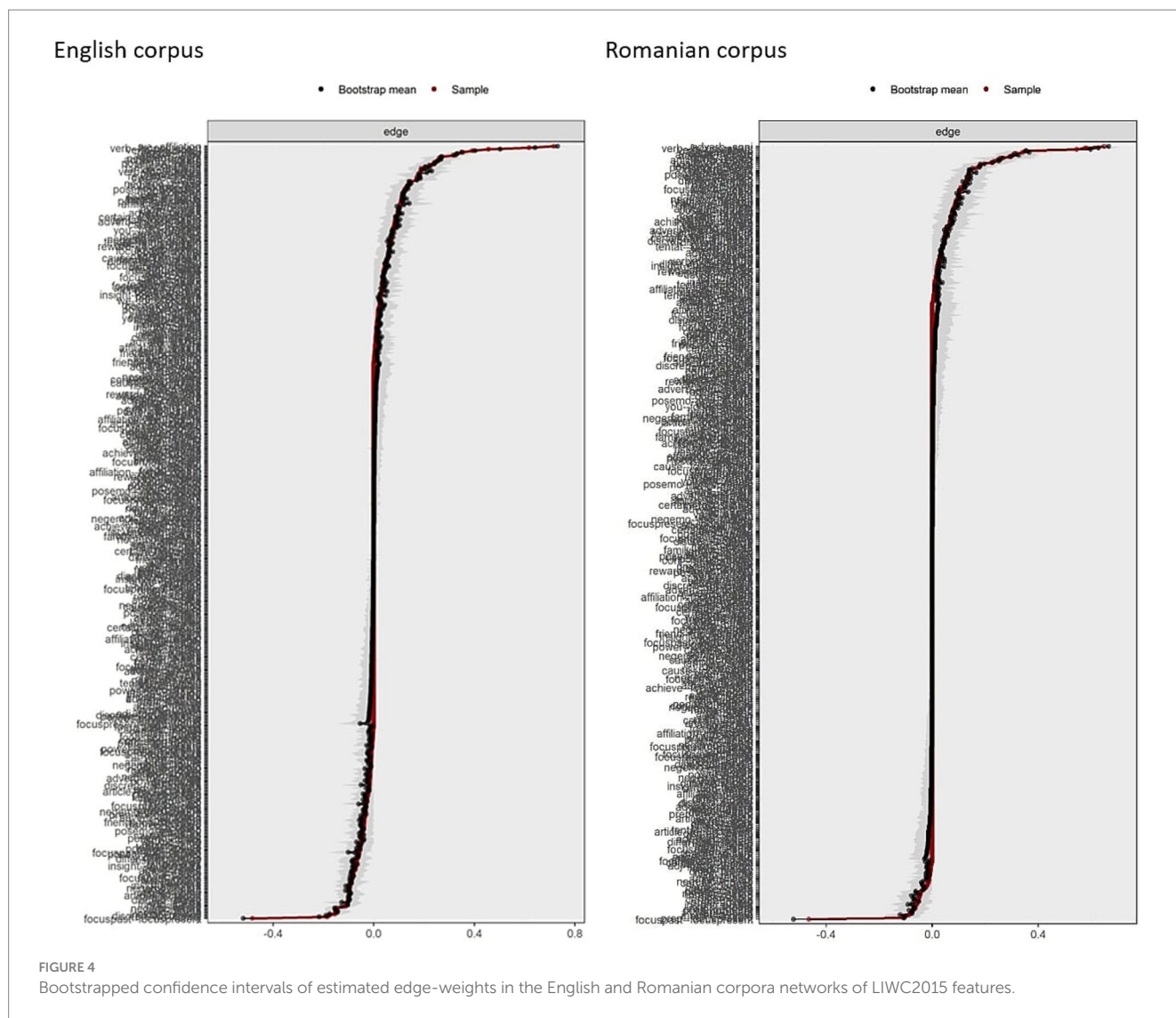
FIGURE 3
Centrality metrics of nodes in the Romanian corpus network.

As inferred from Table 4, distinct linguistic profiles emerged for the clusters in each language when comparing their centroids. A clear divide is present, especially between the two English clusters, one featuring more personal and emotionally expressive language and the other reflecting a more formal, structured style. Interestingly, the Romanian corpus exhibited a third cluster, which could reflect the more diverse academic writing styles in Romanian student texts.

In the English corpus, Cluster 1 was characterized by higher frequencies of personal pronouns and more words involving positive emotions, motivational drives, cognitive processes, and a preoccupation with work, money, home, and friendship. Moreover, this cluster was marked by the use of more verbs and a focus on the present and future, potentially indicating a more action- or goal-oriented approach. All these features, along with more adverbs, conjunctions, and adjectives, suggest a more expressive writing style potentially reflective of less formal academic texts and more personal engagement. Cluster 2 is the opposite of Cluster 1, scoring higher in categories such as articles, prepositions, and focus on the past, with lower use of personal pronouns and emotionally charged language. This indicates a more structured, formal writing style focusing on objective analysis and academic formality, even on topics such as family, religion, or death, which tend to be more specific to humanities or social sciences.

Regarding the patterns that emerged from the Romanian corpus, Cluster 1 stood out for higher frequencies in function words such as personal pronouns, articles, conjunctions, and verbs, as well as in linguistic markers of psychological complexity, as suggested by language referring to emotions (especially negative valence), cognitive processes (with a notable accent on discrepancies and differences), religion, and death. A preoccupation with leisure activities and family matters was also noticed. Thus, overall, the Romanian Cluster 1 was characterized by an emotionally expressive writing style. Cluster 2 was distinct due to its higher scores in LIWC2015 categories like positive emotions, insight, causation, achievement, power, reward, and risk. Moreover, it strongly focused on the future, work, and money, reflecting achievement-oriented or entrepreneurial themes, possibly denoting formal academic texts commonly found in social sciences such as economics, political science, or psychology. Finally, Cluster 3 showed a more disengaged and impersonal profile, with relatively low scores across categories, indicating a less distinctive, more moderate linguistic style that could represent general-purpose or mid-level academic writing.

The Chi-squared test results revealed significant associations between clusters and both genre and discipline in both English ($\chi^2 = 156.46$, $df = 1$, $p < 0.001$ for genre; $\chi^2 = 184.96$, $df = 2$, $p < 0.001$ for discipline) and Romanian corpora ($\chi^2 = 175.11$, $df = 2$, $p < 0.001$ for genre; $\chi^2 = 295.73$, $df = 4$, $p < 0.001$ for discipline). These results



suggest that the linguistic patterns captured by the clustering process might be systematically related to the texts' genre and academic discipline.

By examining the contingency statistics in Table 5 and the linguistic profiles discussed earlier in this subsection, it became apparent that the clusters emerged at the intersection between discipline and genre. This observation could further suggest that specific academic contexts or tasks might require distinct linguistic styles and contents, while disciplines might involve some internal variation in their approaches. For instance, the more personal and expressive cluster derived from the English corpus contained much fewer research and academic papers than the more formal cluster and, consistently, a high percentage of texts from social sciences (about 60% of the texts in Cluster 1). Similarly, many of the English STEM texts (44.8%) were distributed in Cluster 1, which aligns with the fact that a high proportion of papers within this discipline (65.5%) represented coursework and analytical writing. In the Romanian corpus, the three clusters reflected an even more diverse academic writing style. Cluster 1, characterized by emotionally expressive language, was linked to coursework and analytical writing and

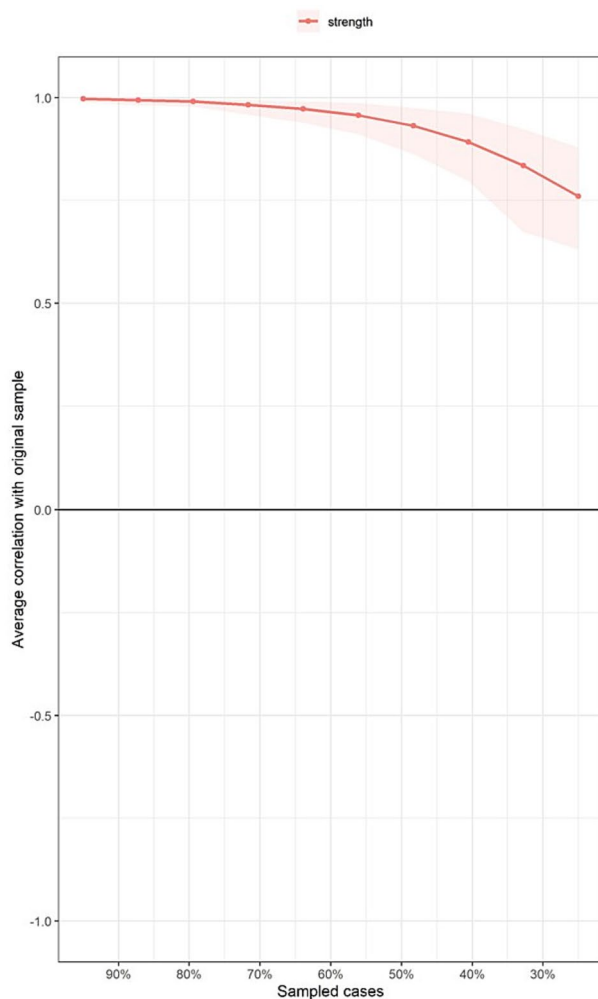
included a high proportion of humanities texts. In contrast, Cluster 2, marked by future orientation and achievement, aligned with the particularities of research papers and social sciences. Cluster 3 represented a general, more detached academic writing style, with very low representation from social sciences and a high concentration of humanities texts, suggesting a second type of coursework and analytical writing within this discipline. Additionally, a relatively high proportion of Romanian research and academic papers (33.5%) were also present in the disengaged-profile Cluster 3, further illustrating the complexity of writing styles within this corpus.

5 Discussion

5.1 Methodological novelty and linguistic insights

This study aimed to elucidate the emotional persona in university students' academic discourse using LIWC2015, a powerful yet easy-to-use tool for automatic language analysis. Specifically, we sought to

English corpus



Romanian corpus

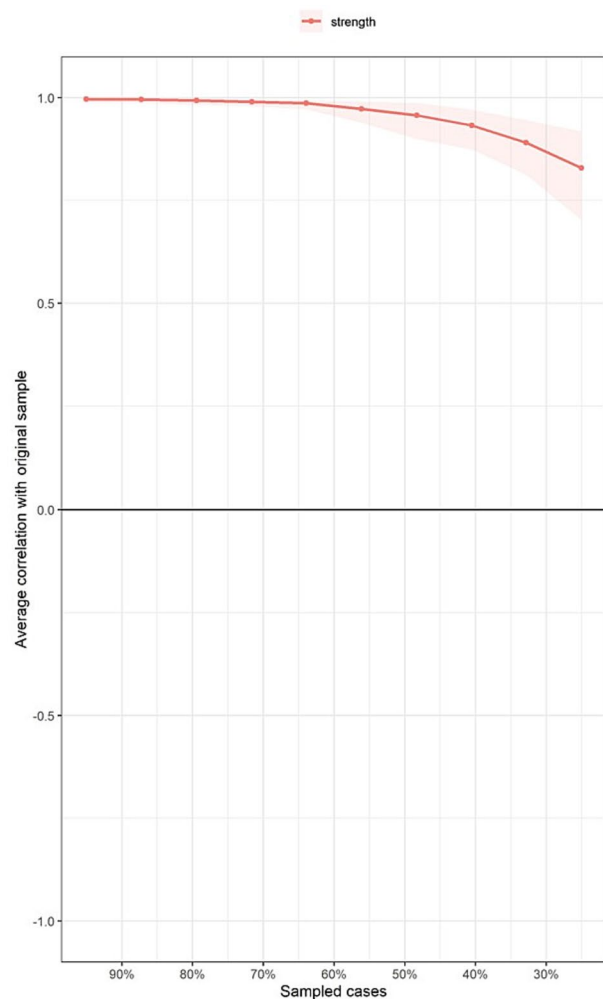


FIGURE 5

Case-dropping bootstrap results for strength centrality in the English and Romanian networks.

explore linguistic patterns across languages, genres, and disciplines, focusing on how multiple linguistic markers varied between English and Romanian academic writing. To this end, we relied on the ROGER corpus, which allowed us to contribute to a research niche regarding Romanian students. This is particularly interesting because Romania, as a former communist country, has undergone numerous socio-economic, cultural, and educational changes in the 35 years since the collapse of the communist regime.

A notable strength of this study lies in its reliance on a corpus collected entirely in Romania, ensuring that the differences observed between Romanian and English academic writing stem from participants within the same demographic and cultural context. Unlike comparative studies that examine texts produced in different countries, this research design allows for a more focused examination of how native versus second-language use interacts with cultural and psychological dimensions. In other words, this framework provides a unique opportunity to explore how linguistic choices in a second language (L2) may favor communication patterns characteristic of the target culture, such as the more direct and personalized style often associated with English, a language rooted in individualistic cultural norms.

Moreover, to gain a deeper understanding of the linguistic markers of academic writing, we implemented a complex data analysis strategy based on machine learning (supervised and unsupervised) and advanced statistical methods. In particular, the network analysis approach for examining how LIWC2015 variables were interconnected was an original choice that could also be valuable in other research contexts that involve this language analysis tool. Thus, our results might offer novel insights into how student writing reflects broader socio-cultural and academic conventions and how the specifics of academic language could be useful in academic writing pedagogy.

The very different classification accuracies of the logistic regression and random forest models highlighted the complexities involved in distinguishing between English and Romanian languages in academic writing based on LIWC2015 features. The logistic regression model performed poorly, indicating that if any linguistic differences between the two languages existed, they could not be captured well by a linear model. The random forest model, by contrast, achieved nearly perfect classification accuracy, suggesting that, indeed, the differences between the English and Romanian texts

TABLE 4 Centroids for LIWC2015 categories by cluster in English and Romanian corpora.

Linguistic features	English corpus		Romanian corpus		
	Cluster 1 <i>n</i> = 499	Cluster 2 <i>n</i> = 621	Cluster 1 <i>n</i> = 313	Cluster 2 <i>n</i> = 145	Cluster 3 <i>n</i> = 349
Function words					
I	0.38	−0.30	0.36	−0.18	−0.25
We	0.41	−0.33	0.19	0.07	−0.2
You	0.40	−0.32	0.14	−0.1	−0.08
Articles	−0.51	0.41	0.22	0.02	−0.21
Prepositions	−0.29	0.23	−0.48	0.39	0.27
Adverbs	0.28	−0.23	0.52	−0.26	−0.36
Conjunctions	0.22	−0.18	0.44	−0.14	−0.34
Verbs	0.60	−0.48	0.6	−0.08	−0.51
Adjectives	0.04	−0.03	−0.01	0.26	−0.1
Affect					
Positive emotions	0.40	−0.32	0.22	0.93	−0.58
Negative emotions	−0.15	0.12	0.5	−0.19	−0.37
Social domain					
Family	−0.21	0.17	0.15	−0.12	−0.08
Friend	0.19	−0.15	0.19	0.25	−0.27
Cognitive processes					
Insight	0.08	−0.06	−0.1	0.27	−0.02
Cause	0.15	−0.12	−0.26	1.02	−0.19
Discrepancy	0.56	−0.45	0.42	0.23	−0.47
Tentative	0.37	−0.30	0.23	0.24	−0.3
Certainty	0.35	−0.28	0.2	0.18	−0.25
Difference	0.29	−0.23	0.29	−0.42	−0.08
Drives					
Affiliation	0.37	−0.30	0.22	0.24	−0.29
Achievement	0.31	−0.25	−0.25	1.51	−0.4
Power	0.02	−0.02	0.12	0.78	−0.43
Reward	0.40	−0.32	−0.13	1.17	−0.37
Risk	0.19	−0.15	0.1	0.51	−0.3
Time orientation					
Focus on the past	−0.3	0.24	0.17	0.01	−0.16
Focus on the present	0.68	−0.54	0.49	−0.1	−0.4
Focus on the future	0.29	−0.24	0.18	0.32	−0.29
Personal concerns					
Work	0.31	−0.25	−0.39	1.26	−0.18
Leisure	−0.15	0.12	0.4	−0.22	−0.27
Home	0.03	−0.03	0.04	0.27	−0.15
Money	0.30	−0.24	−0.29	1.26	−0.26
Religion	−0.24	0.19	0.22	−0.22	−0.11
Death	−0.30	0.24	0.47	−0.36	−0.27

The table contains mean z-scores, which are not suitable for direct interpretations in terms of word percentages. English corpus contained 1,120 texts, while the Romanian corpus contained 807 texts.

TABLE 5 Cluster distribution by discipline and genre in English and Romanian corpora.

	English corpus		Romanian corpus		
	Cluster 1 <i>n</i> = 499	Cluster 2 <i>n</i> = 621	Cluster 1 <i>n</i> = 313	Cluster 2 <i>n</i> = 145	Cluster 3 <i>n</i> = 349
Discipline					
Humanities	34	244	249	26	281
Social sciences	300	174	61	117	36
STEM	165	203	3	2	32
Genre					
Coursework and analytical writing	440	330	285	54	289
Research and academic papers	59	291	28	91	60

English corpus contained 1,120 texts, while the Romanian corpus contained 807 texts.

were significant, though subtle and non-linear or multi-dimensional. The top distinguishing LIWC2015 features in the random forest model – *death, home, family, religion, I, friend, we, you, money, and leisure*—suggested that the expression of personal engagement and writing about important life themes might differ between the two languages. The tendency toward fewer first-person pronouns in Romanian texts may indicate that students adopt a more formal, impersonal tone in academic writing in Romanian. Such a pattern could reflect differing cultural or educational expectations regarding academic discourse, where Romanian academic traditions might emphasize objectivity and detachment. In contrast, English academic writing may prompt more personal involvement and expression.

However, the lack of equivalence in discipline and genre representation across the two languages introduces an additional layer of complexity when comparing the two corpora. Specifically, the Romanian corpus contained a significantly higher proportion of humanities texts, whereas the English corpus included more contributions from STEM. This imbalance could have inadvertently diminished the observed linguistic differences, as the presumably more detached and impersonal nature of STEM writing in English and the more personal and emotionally expressive tone expectable from Romanian humanities writing may have diluted the formal and restrained style often associated with Romanian academic writing. Moreover, cross-linguistic differences might also have been underestimated due to the *foreign language effect*. Research suggests that thinking and writing in a non-native language can reduce the influence of emotions and encourage more logical, rational thinking (Circi et al., 2021; Hayakawa et al., 2022; Keysar et al., 2012). Thus, writing in English (L2) might require heightened cognitive control, leading to simplification or a shift toward rationality over emotional depth. However, it is noteworthy that the results did not reveal clear patterns of higher emotionality in the Romanian (L1) texts, suggesting that other psychological, cultural, or contextual factors may play a role in shaping the emotional personas when writing in these languages.

The network analysis approach revealed a distinction in the interconnectedness of linguistic features between the English and Romanian corpora, with the English corpus displaying higher overall connectedness, as evidenced by the greater number of edges. This might suggest that students tended to integrate various linguistic elements more cohesively when writing in English as a second language, potentially reflecting their adaptation to the linear, argument-driven structure typical of English-language academic

discourse (Hinkel, 2002). Nevertheless, the centrality metrics showed that while both languages emphasize action-oriented and motivational language, the English network illustrated a more personal and self-reflective tone. In other words, Romanian writing remained more formal and detached, which aligns with previous research (Bercuci and Chitez, 2023).

The cluster analysis revealed two linguistic profiles within the English corpus and three within the Romanian corpus. A third cluster in the Romanian corpus could underscore more diverse academic writing in this linguistic context, which might reflect the transitional state of Romanian academic writing, where traditional genres and styles coexist with more contemporary, global academic conventions, as emphasized, for instance, by Băniceru et al. (2012) and Chitez and Kruse (2012). The clusters derived from the linguistic features were significantly associated with both genre and discipline. In the English corpus, Cluster 1, characterized by a more personal and expressive style, was predominantly composed of coursework and analytical writing, and social sciences texts. Cluster 2, which exhibited a more formal and structured style, was more heavily associated with research papers. Similarly, in the Romanian corpus, Cluster 1 contained more emotionally expressive language and was strongly linked to coursework and analytical writing, while Cluster 2, with its focus on achievement and future orientation, was more common in social sciences and research papers. Cluster 3 was characterized by a general, detached academic writing style, with a substantial concentration of humanities texts, a minimal representation from social sciences, and the inclusion of nearly all STEM texts.

This study represents a novel exploration of the emotional dimensions of Romanian academic writing, a field that has been largely neglected in prior research. By using the bilingual ROGER corpus, the first to comprehensively capture the state of university academic writing in Romania, we offer unique insights into the interplay between language, emotion, and academic conventions. The original dataset enables a bilingual comparative approach that highlights cross-linguistic differences and cultural nuances in academic discourse. Furthermore, the methodological approach employed in this research, i.e., integrating LIWC for automated emotional analysis, represents a groundbreaking advancement in Romanian academic writing studies. Unlike previous research, which focused primarily on structural or rhetorical features, this study introduces the psychological dimension by capturing the emotional persona embedded in student writing. By uncovering significant

contrasts in emotional expression across languages and disciplines, our study not only enriches the understanding of Romanian academic discourse but also contributes to the broader field of multilingual academic writing. The analysis offers critical insights for developing culturally sensitive teaching methodologies that address the emotional and linguistic needs of students navigating multilingual academic environments.

5.2 Culturally shaped linguistic features of Romanian academic writing

The emotional persona of Romanian students' academic writing emerges as a distinctive interplay of linguistic markers shaped by cultural, disciplinary, and contextual influences. Our study identifies several key features that distinguish Romanian students' academic writing in L1 from academic writing in L2 English, drawing on the analysis of 33 LIWC2015 features and the structural relationships among these features.

As David's (2015) comprehensive study on the psychological profile of Romanians highlights, the culture is characterized by a blend of emotional restraint, collectivism, and a tendency toward skepticism and indirect communication. These traits are deeply embedded in Romanian social norms, influencing academic traditions and communication styles.

As explained below, this profile is mirrored in students' academic writing in Romanian. However, interestingly, when writing in English, students tend to express a different style, as though they adopt, at least partly, a different academic persona that struggles to depart from the Romanian-specific restraint style and embrace a more personal, direct, and expressive communication, which is in line with Anglo-Saxon conventions. This difference might illustrate the tension between culturally and educationally ingrained communication norms and the need to adapt to global academic standards.

Specifically, the analysis presented in the current study reveals that Romanian academic writing is characterized by a more formal, detached style, as evidenced by the significantly lower use of first-person pronouns compared to English texts. This trend reflects broader cultural norms in Romania, where academic traditions emphasize objectivity and deference to authority over personal engagement. In contrast, academic writing in L2 English displays greater use of personal pronouns and emotionally expressive language, indicating a shift toward the assertive and individualistic norms of Anglo-Saxon academic conventions. The random forest model results highlight the importance of linguistic markers related to personal concerns, such as "death," "home," and "family," as well as personal pronouns like "I" and "we." While these features are prominent in distinguishing between English and Romanian texts, their relative frequencies suggest a nuanced linguistic style in Romanian academic writing. For instance, Romanian texts often avoided direct references to the self, aligning with the cultural emphasis on collective expression and indirect communication. Network analyses further reveal distinct patterns of interaction among linguistic features in Romanian academic texts. Compared to L2 English writing, the Romanian corpus exhibited fewer connections between linguistic variables, indicating a less cohesive integration of elements. Key features such as "discrepancy," "positive emotions," and "focus on the present" emerge as central in shaping the structure of Romanian texts. These features

serve as bridges, connecting otherwise disparate linguistic markers and facilitating the transition between ideas. This indicates that Romanian students rely on nuanced language to maintain flow and coherence in their writing, despite a generally formal and restrained emotional tone.

The cluster analysis provides additional insights into the diversity of writing styles in Romanian versus L2 English. In the English L2 corpus, two distinct clusters emerge: one characterized by a personal and expressive style and the other by a more formal and structured approach. The expressive cluster features a higher use of personal pronouns, positive emotion words, and markers of motivational drives, reflecting a goal-oriented and engaging tone. This style, often found in coursework and analytical writing, aligns with Anglo-Saxon academic norms that encourage individual expression and critical engagement. Conversely, the formal cluster, associated with research papers, is marked by higher frequencies of articles, prepositions, and a focus on past events, indicative of objective analysis and academic rigor.

In comparison, the Romanian corpus exhibits three clusters, highlighting greater diversity in writing styles. The first cluster, marked by emotionally expressive language, shares similarities with the English expressive cluster but includes a notable emphasis on negative emotional markers and cognitive processes such as "discrepancy" and "difference." This suggests a more reflective and complex emotional engagement, particularly in less formal academic contexts like coursework. The second cluster, distinguished by future orientation and markers of achievement and power, aligns with the English formal cluster but exhibits stronger motivational themes, likely reflecting the influence of social sciences and research-oriented writing. The third Romanian cluster represents a detached and impersonal style, with low frequencies across most linguistic categories, reflecting a neutral tone often associated with general-purpose academic writing.

Such distinctions underscore the influence of cultural norms on academic writing. While English texts often reflect a balance between expressiveness and structure, Romanian texts exhibit a stronger separation between emotional engagement and formal academic norms. The additional cluster in the Romanian corpus suggests a transitional stage, where traditional academic expectations coexist with emerging global influences, creating a broader spectrum of styles. The comparison highlights the challenges faced by Romanian students as they adapt to bilingual academic expectations. The expressive styles in both corpora indicate a shift toward greater emotional engagement in less formal contexts, while the formal styles reflect ongoing adherence to disciplinary conventions. By understanding these patterns, educators can better support students in navigating the linguistic and cultural complexities of multilingual academic writing.

5.3 Pedagogical implications

Study findings offer several key takeaways for teaching practices at the university level, particularly in multilingual and multicultural academic settings. First, the distinct linguistic profiles identified in Romanian academic writing, ranging from formal and detached styles to emotionally expressive approaches, highlight the need for pedagogical strategies that address this diversity. Educators should recognize and accommodate the influence of cultural norms on

writing, especially the preference for objectivity and formality in Romanian academic traditions. Tailored instruction can help students balance these norms with the more personal and assertive styles encouraged in English academic writing.

Second, the contrast between the cohesive, highly connected linguistic networks in English texts and the more segmented structure of Romanian writing suggests a need for targeted training in integrating linguistic elements cohesively. Workshops focusing on the use of connectors, cohesive devices, and argumentation strategies could bridge this gap, helping students produce writing that aligns with global academic expectations while maintaining their unique cultural perspective.

Finally, the findings on emotional personas in writing provide an opportunity to incorporate discussions of voice, tone, and audience into writing curricula. By encouraging students to explore how emotional engagement enhances clarity and persuasiveness in their texts, educators can encourage greater confidence in navigating different academic conventions. Addressing these issues explicitly in coursework could enable students to adapt their writing more effectively across genres, disciplines, and cultural contexts.

In sum, our research highlights the importance of a nuanced, culturally informed approach to teaching academic writing at the university level. By leveraging these insights, educators can support Romanian students in developing versatile, internationally competitive writing skills while respecting and integrating their linguistic and cultural heritage. This dual emphasis ensures that students are not only prepared to meet international academic standards but are also empowered to contribute their unique voices to the broader academic conversation.

5.4 Study limitations and prospects for future research

Given the dual challenges posed by Romania's socio-historical context and the demands of multilingual academic writing, the current study sought to build on a critical gap in understanding how emotional personas are reflected in student writing. However, the topic of emotional persona in academic writing is complex, and as with any study, our research is not without its limitations, which present opportunities for further exploration and development.

First, certain methodological shortcomings warrant further exploration. In this regard, the data was limited to a sample of Romanian students from nine universities – all state institutions – which may affect the generalizability of the findings to other cultural or linguistic contexts or even to the population of Romanian university students. The ROGER corpus already offers broad coverage, but the sample could not be considered nationally representative. Similarly, our dataset was marked by several class imbalances, which could have introduced a confounding effect in all our findings, especially those related to network and cluster analysis. Moreover, the recoding process of the genre and discipline variables did not involve multiple raters or a rigorous methodology, which could have impacted the quality of the new classes.

To address these methodological challenges, replication studies and efforts to refine variable control are needed. Moreover, future studies could expand the scope to include students from different linguistic

backgrounds, allowing for a more comprehensive understanding of how emotional personas are expressed across different languages and academic traditions. LIWC2015 could provide the technical means to extend the current research to a multilingual, intercultural context, revealing valuable insights into the emotional and cognitive aspects of academic writing. However, its closed-vocabulary nature may overlook some of the more nuanced or context-specific elements of student writing. Future research could also explore how open-vocabulary approaches, which allow for analyzing emergent linguistic patterns, could complement the findings based on LIWC dictionaries, which follow a list of predefined linguistic features.

Second, the main goal of the current study was to understand whether different emotional personas are present in Romanian versus English academic writing and whether discipline- and genre-specific linguistic patterns exist. This research topic, while valuable, opens the door to numerous related questions. For instance, while our study focused on Romanian students' one-time written academic discourse, future research could explore how emotional personas evolve over time. Thus, a longitudinal perspective could bring a deeper understanding of how academic writing skills – and the emotional personas embedded within them – develop as students advance through their academic careers. An additional valuable question would be whether tailored pedagogical approaches could help students refine their emotional personas in academic writing and whether such refinements could foster improved communication, critical thinking skills, motivation, or cultural adaptability.

6 Conclusion

This study sheds light on the emotional and cognitive characteristics of Romanian (L1) and English (L2) student writing, revealing significant cross-linguistic, as well as discipline- and genre-specific patterns. By leveraging the LIWC2015 tool alongside machine learning and network analysis, we identified distinct linguistic profiles in the Romanian and English corpora. These results might suggest the role of the second language (L2) cultural norms in shaping academic writing and emotional expression. Our findings contribute to a deeper understanding of the complex interplay between psycholinguistic and cultural factors, offering valuable insights for educators and researchers in multilingual academic settings.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

DD: Writing – original draft, Formal analysis, Methodology, Validation, Visualization, Writing – review & editing. MC: Writing – original draft, Conceptualization, Resources, Data curation, Writing – review & editing. FS: Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This publication was funded by the West University of Timișoara through the research fund allocated to Romanian public higher education institutions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Akobeng, A. K. (2007). Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatr.* 96, 644–647. doi: 10.1111/j.1651-2227.2006.00178.x
- Alves, R. A., and Haas, C. (2012). Special issue: writing and cognition, in honor of John R. Hayes: editors' introduction. *Writ. Commun.* 29, 239–243. doi: 10.1177/0741088312453135
- Argaman, O. (2010). Linguistic markers and emotional intensity. *J. Psycholinguist. Res.* 39, 89–99. doi: 10.1007/s10936-009-9127-1
- Băniceru, C., Borchin, M. I., Doroholschi, C., and Tucan, D. (2012). Academic writing in Romania: a contrastive analysis of BA thesis introductions in Romanian and English. *Quaestiones Roman.* 1, 331–345.
- Bazerman, C., and Paradis, J. (1991). *Textual dynamics of the professions: Historical and contemporary studies of writing in academic and other professional communities*. Madison, WI: University of Wisconsin Press.
- Bercuci, L., and Chitez, M. (2023). "The influence of the Romanian academic style on student writing in English" in New to the LSP classroom? A selection of monographs on successful practices. ed. M. Vranova (Wilmington, DE: Vernon Press), 169–185.
- Bereiter, C. (1980). "Development in writing" in Cognitive processes in writing. eds. L. W. Gregg and E. R. Steinberg (Hillsdale, NJ: Lawrence Erlbaum), 73–93.
- Bereiter, C., and Scardamalia, M. (2013). *The psychology of written composition*. London: Routledge.
- Bischi, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., et al (2016). mlr: Machine learning in R. *Journal of Machine Learning Research*, 17.
- Boc, O. (2020). The importance of multilingualism in the context of globalization. In International annual symposium of the institute-of-Romanian-philology - a Philippide - Iasi-branch of the Romanian-academy.
- Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., et al. (2021). Network analysis of multivariate data in psychological science. *Nat. Rev. Methods Primers* 1:58. doi: 10.1038/s43586-021-00055-w
- Bowerman, B. L., O'Connell, R. T., and Murphree, E. S. (2015). *Regression analysis: Unified concepts, practical applications, and computer implementation*. New York, NY: Business Expert Press.
- Boyd, R. L., and Schwartz, H. A. (2021). Natural language analysis and the psychology of verbal behavior: The past, present, and future states of the field. *Journal of Language and Social Psychology*, 40, 21–41. doi: 10.1177/0261927X20967028
- Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Bromberek-Dyzman, K., Jończyk, R., Vasileanu, M., Niculescu-Gorpin, A. G., and Băk, H. (2021). Cross-linguistic differences affect emotion and emotion-laden word processing: evidence from Polish-English and Romanian-English bilinguals. *Int. J. Biling.* 25, 1161–1182. doi: 10.1177/1367006920987306
- Bucur, A. M., Chitez, M., Mureșan, V., Dincă, A., and Rogobete, R. (2022). Ro-AWL: building the first Romanian academic word list through corpus analysis. In Proceedings of the thirteenth language resources and evaluation conference (pp. 4739–4746).
- Cameron, J., Nairn, K., and Higgins, J. (2009). Demystifying academic writing: reflections on emotions, know-how and academic identity. *J. Geogr. High. Educ.* 33, 269–284. doi: 10.1080/03098260902734943
- Carless, D., and Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assess. Eval. High. Educ.* 43, 1315–1325. doi: 10.1080/02602938.2018.1463354
- Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *J. Stat. Softw.* 61, 1–36. doi: 10.18637/jss.v061.i06
- Chen, S. (2024). Mining emotions in academic writing: a subdisciplinary probe into medical research articles. *Engl. Specif. Purp.* 74, 55–67. doi: 10.1016/j.esp.2024.01.002
- Cheng, T. Y. (1993). Is Chinese thought pattern linear, circular or both? Available at: <https://eric.ed.gov/?id=ED364098> (Accessed October 18, 2024).
- Chitez, M., Bercuci, L., Dincă, A., Rogobete, R., and Csűrös, K. (2021). *Corpus of Romanian academic genres (ROGER)*. Timișoara, Romania: West University of Timisoara.
- Chitez, M., and Kruse, O. (2012). Writing cultures and genres in European higher education. In *University writing: Selves and texts in academic societies*. (Eds.), M. Castelló and C. Donahue, Leiden, Netherlands: Brill. (pp. 151–175). doi: 10.1163/9781780523873_010
- Circi, R., Gatti, D., Russo, V., and Vecchi, T. (2021). The foreign language effect on decision-making: a meta-analysis. *Psychon. Bull. Rev.* 28, 1131–1141. doi: 10.3758/s13423-020-01871-z
- Cojocaru, V. (2020). English discourse markers in spoken Romanian: pragmatic borrowings or a code-switching phenomenon? *Rev. Roumaine Linguist.* 65, 261–271.
- Cojocaru, V. (2021). The use of discourse markers by foreign students learning Romanian. *Rev. Roumaine Linguist.* 66, 335–343.
- Connor, U. (1996). *Contrastive rhetoric*. Cambridge, UK: Cambridge University Press.
- David, D. (2015). *Psihologia poporului român: Profilul psihologic al românilor într-o monografie cognitiv-experimentală* [the psychology of the Romanian people: the psychological profile of Romanians in a cognitive-experimental monograph] Iași, Romania: Polirom.
- Deserno, M. K., Isvoranu, A. M., Epskamp, S., and Blanken, T. F. (2022). "Descriptive analysis of network structures in Network psychometrics with R. A guide for behavioral and social scientists." eds. A.-M. Isvoranu, S. Epskamp, L. Waldorp and D. Borsboom. (London: Routledge), 45–65.
- Dincă, A., Bucur, A.-M., and Chitez, M. (2024). Tracking phraseological inter-complexity: a corpus-based contrastive analysis of learner English and Romanian novice academic writing. Paper presented at the Learner Corpus Research Conference 2024, Tartu, Estonia, September 5–7.
- Doroholschi, C. I., Tucan, D., Chitez, M., and Kruse, O. (2018). "Introduction: understanding academic writing in the context of central and eastern European higher education" in University writing in central and Eastern Europe: Tradition, transition, and innovation. eds. M. Chitez, C. I. Doroholschi, O. Kruse, L. Salski and D. Tucan (Berlin: Springer), 1–12.
- Dudău, D. P., and Sava, F. A. (2021). Performing multilingual analysis with linguistic inquiry and word count 2015 (LIWC2015). An equivalence study of four languages. *Front. Psychol.* 12:570568. doi: 10.3389/fpsyg.2021.570568
- Dudău, D. P., and Sava, F. A. (2022). The development and validation of the Romanian version of linguistic inquiry and word count 2015 (Ro-LIWC2015). *Curr. Psychol.* 41, 3597–3614. doi: 10.1007/s12144-020-00872-4
- Eichstaedt, J. C., Kern, M. L., Yaden, D. B., Schwartz, H. A., Giorgi, S., Park, G., et al. (2021). Closed- and open-vocabulary approaches to text analysis: a review, quantitative comparison, and recommendations. *Psychol. Methods* 26, 398–427. doi: 10.1037/met0000349
- Epskamp, S., Borsboom, D., and Fried, E. I. (2018). Estimating psychological networks and their accuracy: a tutorial paper. *Behav. Res. Methods* 50, 195–212. doi: 10.3758/s13428-017-0862-1
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. (2012). Qgraph: network visualizations of relationships in psychometric data. *J. Stat. Softw.* 48, 1–18. doi: 10.18637/jss.v048.i04

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Epskamp, S., and Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychol. Methods* 23, 617–634. doi: 10.1037/met0000167
- Ferrari, G., and Boca, L. (2017). Metaphors in foreign language teaching: English and Romanian business metaphors. *European Proceedings of Social and Behavioural Sciences*.
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. California: SAGE Publications Ltd.
- Fløttum, K. (2012). “Variation of stance across culture” in *Stance and voice in written academic genres*. eds. K. Hyland and C. Sancho Guinda (Houndsmill: Palgrave MacMillan).
- Ghivirigă, T., and Baci, L. C. (2015). Between epistemic and deontic: the use of modal verbs in a Corpus of Romanian scientific texts on economics. In *International multidisciplinary scientific conference on social sciences and arts SGEM 2015* (pp. 105–112).
- Gogăță, C. (2023). Food metaphors-A form of cultural heritage and linguistic innovation. A 1068 comparison between Romanian, English, and Italian food metaphors. *Studia Universitatis Babeş-Bolyai Philologia*, 68, 121–140. doi: 10.24193/subbphilo.2023.1.07
- Hairston, M. (1982). The winds of change: Thomas Kuhn and the revolution in the teaching of writing. *College Compos. Commun* 33, 76–88. doi: 10.58680/cc198215868
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning. Data mining, inference, and prediction*. 2nd Edn. New York, NY: Springer Science + Business Media. doi: 10.1007/978-0-387-84858-7
- Hayakawa, S., Pan, Y., and Marian, V. (2022). Language changes medical judgments and beliefs. *Int. J. Biling.* 26, 104–121. doi: 10.1177/13670069211022851
- Hayes, J. R., and Flower, L. S. (1980). Writing as problem solving. *Visible Language*, 14, 388–399.
- Hinkel, E. (2001). Matters of cohesion in L2 academic texts. *Appl. Lang. Learn.* 12, 111–132.
- Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*. London: Routledge.
- Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Appl. Linguis.* 17, 433–454. doi: 10.1093/applin/17.4.433
- Hyland, K. (1998). Hedging in scientific research articles. Amsterdam, Netherlands: John Benjamins.
- Hyland, K. (1999). Academic attribution: citation and the construction of disciplinary knowledge. *Appl. Linguis.* 20, 341–367. doi: 10.1093/applin/20.3.341
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London, UK: Longman.
- Hyland, K. (2002). Authority and invisibility: authorial identity in academic writing. *J. Pragmat.* 34, 1091–1112. doi: 10.1016/S0378-2166(02)00035-8
- Hyland, K. (2004). *Disciplinary discourses, Michigan classics ed.: Social interactions in academic writing*. Ann Arbor, MI: University of Michigan Press.
- Hyland, K. (2005). *Metadiscourse: exploring interaction in writing*. London, UK: Continuum.
- Hyland, K. (2008). As can be seen: lexical bundles and disciplinary variation. *Engl. Specif. Purp.* 27, 4–21. doi: 10.1016/j.esp.2007.06.001
- Hyland, K. (2009). *Academic discourse: English in a global context*. London, UK: Continuum.
- Hyland, K. (2012). *Disciplinary identities: Individuality and community in academic writing*. Cambridge: Cambridge University Press.
- Hyland, K., and Tse, P. (2007). Is there an “academic vocabulary”? *TESOL Q.* 41, 235–253. doi: 10.1002/j.1545-7249.2007.tb00058.x
- Ivanič, R. (1998). Writing and identity: The discursive construction of identity in academic writing. Amsterdam, Netherlands: John Benjamins. doi: 10.1075/swll.5
- Kahn, J. H., Tobin, R. M., Massey, A. E., and Anderson, J. A. (2007). Measuring emotional expression with the linguistic inquiry and word count. *Am. J. Psychol.* 120, 263–286. doi: 10.2307/20445398
- Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural education. *Lang. Learn.* 16, 1–20. doi: 10.1111/j.1467-1770.1966.tb00804.x
- Keysar, B., Hayakawa, S. L., and An, S. G. (2012). The foreign-language effect: thinking in a foreign tongue reduces decision biases. *Psychol. Sci.* 23, 661–668. doi: 10.1177/0956797611432178
- Kruse, O., Chitez, M., Rodriguez, B., and Castelló, M. (Eds.). (2016). *Exploring European writing cultures: Country reports on genres, writing practices, and languages used in European higher education*. (Working Papers in Applied Linguistics No. 10) Winterthur, Switzerland: ZHAW Zurich. University of Applied Sciences.
- Kučera, D., and Mehl, M. R. (2022). Beyond English: Considering language and culture in psychological text analysis. *Front. Psychol.* 13:819543. doi: 10.3389/fpsyg.2022.819543
- Langer, J. A., and Applebee, A. N. (1987). *How writing shapes thinking: A study of teaching and learning*. Urbana, IL: National Council of Teachers of English.
- Lesmeister, C. (2015). *Mastering machine learning with R*. Birmingham, UK: Packt Publishing Ltd.
- Liu, X. (2013). Evaluation in Chinese university EFL students' English argumentative writing: An APPRAISAL study. *Electr. J. Foreign Lang. Teach.* 10, 40–53.
- Machin, D., and Mayr, A. (2012). *How to do critical discourse analysis: A multimodal introduction*. London, UK: SAGE Publications.
- Manda, I., and Chitez, M. (2022). Writing culture clash: English essay writing challenges in Romanian pre-university settings. In *Conference proceedings of the future of education 2022* (pp. 403–409).
- Meier, T., Boyd, R. L., Pennebaker, J. W., Mehl, M. R., Martin, M., Wolf, M., et al. (2018). “LIWC auf Deutsch”: the development, psychometrics, and introduction of DE-LIWC2015. Available at: <https://osf.io/tfqzc/> (Accessed October 12, 2024).
- Moffett, J. (1968). *Teaching the universe of discourse*. Boston, MA: Houghton Mifflin.
- Monroe, J. (Ed.) (2002). *Writing and revising in the disciplines*. Ithaca, NY: Cornell University Press.
- Morita, N. (2004). Negotiating participation and identity in second language academic communities. *TESOL Q.* 38, 573–603. doi: 10.2307/3588281
- Muresan, V., Rogobete, R., Bucur, A. M., Chitez, M., and Dincă, A. (2022). “Phraseology in Romanian academic writing: Corpus-based explorations into field-specific multiword units” in *Recent advances in digital humanities: Romance language applications*. eds. D. Anca, M. Chitez, L. Dinu and M. Dobre (Berlin, Germany: Peter Lang), 29–48.
- Murray, D. M. (1968). *A writer teaches writing: A practical method of teaching composition*. Boston, MA: Houghton Mifflin.
- Negri, A., Andreoli, G., Barazzetti, A., Zamin, C., and Christian, C. (2020). Linguistic markers of the emotion elaboration surrounding the confinement period in the Italian epicenter of COVID-19 outbreak. *Front. Psychol.* 11:568281. doi: 10.3389/fpsyg.2020.568281
- Neuendorf, K. A. (2017). *The content analysis guidebook*. 2nd Edn. Thousand Oaks, CA: SAGE Publications.
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* 41, 673–690. doi: 10.1007/s11135-006-9018-6
- Oravița, A., Chitez, M., Bercuci, L., and Rogobete, R. (2022). “Using the bilingual Corpus of Romanian academic genres (ROGER) platform to improve students' academic writing” in *Intelligent CALL, granular systems, and learner data: Short papers from EUROCALL 2022*. eds. B. Arnbjörnsdóttir, B. Bédi, L. Bradley, K. Friðriksdóttir, H. Garðarsdóttir and S. Thouéšny, et al. Research-publishing.net, 315–321. doi: 10.14705/rpnet.2022.61.1477
- Pennebaker, J. W., Boyd, R. L., Jordan, K., and Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: University of Texas at Austin.
- Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., and Beaver, D. I. (2014). When small words foretell academic success: the case of college admissions essays. *PLoS One* 9:e115844. doi: 10.1371/journal.pone.0115844
- Poe, M., Lerner, N., and Craig, J. (2010). *Learning to communicate in science and engineering: Case studies from MIT*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/7702.001.0001
- Pohl, T. (2007). *Studien zur Ontogenese wissenschaftlichen Schreibens*. Tübingen, Germany: Max Niemeyer.
- Popescu, T. (2017). Culture-Bo und metaphors. A corpus-based comparative approach to English and Romanian journals. *Eur. Proc. Soc. Behav. Sci.* doi: 10.15405/epsbs.2017.05.02.128
- Rogoff, B. (2003). *The cultural nature of human development*. New York, NY: Oxford University Press.
- Senar, F., Janés, J., Serrat, E., and Huguet, Á. (2024). Assessing the relationship between L1 knowledge and fluid intelligence in second language acquisition: the case of immigrant students in Catalonia. *Int. J. Biling. Educ. Biling.* 27, 1008–1020. doi: 10.1080/13670050.2024.2320310
- Silvano, P., Damova, M., Valūnaitė Oleškevičienė, G., Liebeskind, C., Chiarcos, C., Trajanov, D., et al. (2022). IO-based annotated multilingual parallel corpus for discourse markers. In *Language Resources and Evaluation Conference, LREC 2022*. (Eds.), N. Calzolari, F. Bechet, P. Blache, K. Choukri, C. Cieri, T. Declerck, et al. Paris, France: European Language Resources Association. (pp. 2739–2749).
- Strilețchi, C., Chitez, M., and Csűrös, K. (2022). Building Roger: technical challenges while developing a bilingual corpus management and query platform. In *Proceedings of the 17th international conference on software technologies (ICSOFT)*, 10–12 July 2022, Lisbon, Portugal (pp. 390–398).
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge, UK: Cambridge University Press.
- Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.
- Tausczik, Y. R., and Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* 29, 24–54. doi: 10.1177/0261927X09351676
- Thaiss, C., and Zawacki, T. M. (2006). *Engaged writers and dynamic disciplines: Research on the academic writing life*. Portsmouth, NH: Boynton/Cook.

- Trantescu, A. M., and Reiss, G. (2022). Considerations on the meaning and translation of English heart idioms. Integrating the cognitive linguistic approach. *Open. Linguistics* 8, 427–439. doi: 10.1515/opli-2022-0203
- Tucan, D., Pop, A.-M., Bercuci, L., and Chitez, M. (2020). Writing transition: a corpus-supported study of the freshmen's written argumentation competence building. In Conference proceedings of the 13th international conference innovation in language learning (pp. 62–66).
- Tuszynski, J. (2021). caTools: tools: moving window statistics, GIF, Base64, ROC AUC, etc. *R package version 1:2*.
- Varttala, T. (2001). Hedging in scientifically oriented discourse: Exploring variation according to discipline and intended audience. Tampere, Finland: Tampere University Press.
- Vinkers, C. H., Tjebkink, J. K., and Otte, W. M. (2015). Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis. *BMJ* 351:h6467. doi: 10.1136/bmj.h6467
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., et al. (2019). Welcome to the tidyverse. *J. Open Source Soft.* 4:1686. doi: 10.21105/joss.01686
- Williams, B. T. (2017). *Literacy practices and perceptions of agency: Composing identities*. London: Routledge.
- Xiao, W., Guo, Y., and Zhao, X. (2023). Towards positivity: a large-scale diachronic sentiment analysis of the humanities and social sciences in China. *Fudan J. Humanit. Soc. Sci.* 16, 569–589. doi: 10.1007/s40647-023-00380-2



OPEN ACCESS

EDITED BY

Xinghua Liu,
Shanghai Jiao Tong University, China

REVIEWED BY

Md. Aftab Uddin,
University of Chittagong, Bangladesh
Sweety Shah,
I.J. Institute of Management Studies, India

*CORRESPONDENCE

Qingran Wang
✉ cu192019@cupl.edu.cn

RECEIVED 21 October 2024

ACCEPTED 18 December 2024

PUBLISHED 20 January 2025

CITATION

Wang Q (2025) EFL learners' motivation and acceptance of using large language models in English academic writing: an extension of the UTAUT model.
Front. Psychol. 15:1514545.
doi: 10.3389/fpsyg.2024.1514545

COPYRIGHT

© 2025 Wang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

EFL learners' motivation and acceptance of using large language models in English academic writing: an extension of the UTAUT model

Qingran Wang*

School of Foreign Studies, China University of Political Science and Law, Haidian, Beijing, China

Large language models (LLMs), represented by ChatGPT, are one of the most significant technological breakthroughs in generative AI and have begun to be applied in EFL writing instruction. The advent of LLMs presents both opportunities and challenges for EFL learners, underscoring the importance of empirical evidence on their motivation and acceptance of using LLMs in learning English academic writing. This study recruited 238 participants who had completed one semester of training in using LLMs for business-related English academic writing. Participants answered question items based on the L2 Motivational Self System and the Unified Theory of Acceptance and Use of Technology (UTAUT). Partial least squares structural equation modeling (PLS-SEM) was employed to examine the structural relationships between the variables of motivation, region, previous learning experience, and the UTAUT model. Additionally, the moderating effect of motivation on the relationship between the four UTAUT determinants, behavioral intention, and use behavior was tested. The results show that performance expectancy and social influence significantly affect learners' behavioral intention to use LLMs. Moreover, motivation proved to be a key factor in shaping both behavioral intention and actual use behavior, highlighting its crucial role in the adoption of technology for learning English academic writing.

KEYWORDS

large language models (LLMs), academic writing, motivation, unified theory of acceptance and use of technology (UTAUT), EFL courses

1 Introduction

The integration of advanced technologies in education has significantly transformed how students approach learning, particularly in fields like business-related English academic writing (Xu and Wang, 2024a). One notable development is the advent of large language models (LLMs) such as OpenAI's GPT series, which have shown immense potential in assisting learners by generating coherent text, providing real-time feedback, and improving the overall quality of written work (Liu and Ma, 2024; Su et al., 2023; Lu et al., 2024). At the same time, accelerated urbanization in China has fueled rapid economic growth but also exacerbated educational disparities between urban and rural areas (Zhang, 2017). This divide is evident in the varying English proficiency levels of English as a Foreign Language (EFL) learners in Chinese universities (Murray et al., 2023). To address these disparities, UNESCO advocates for the use of artificial intelligence (AI) to promote educational equity by catering to diverse learner needs and fostering inclusive education (Holmes and Miao, 2023). However, while the application of LLMs in English as a Foreign Language (EFL) learning contexts holds great promise, the factors influencing the acceptance and adoption of these tools by Chinese

rural and urban EFL learners in higher education, particularly in academic writing, remain underexplored.

This study addresses this gap by investigating the factors influencing the adoption of LLMs for learning business-related English academic writing among Chinese rural and urban EFL learners. These factors include motivation, regional background, previous learning experience, performance expectancy, effort expectancy, social influence, and facilitating condition, with the Unified Theory of Acceptance and Use of Technology (UTAUT) model proposed by Venkatesh et al. (2003) serving as the theoretical foundation. The UTAUT model is the leading framework used by academia to explore users' acceptance of technology (Wu et al., 2019), considering performance expectancy, effort expectancy, social influence, and facilitating conditions as key predictors of technology acceptance. These variables are particularly useful in understanding how students perceive and utilize digital tools for EFL learning (Guggemos et al., 2020; Hsu, 2023; Menon and Shilpa, 2023; Grassini et al., 2024). To expand the model's explanatory power, motivation is introduced as both a predictor and a moderating variable, while students' region and previous learning experiences (including whether they have ever taken computer courses and their streams in high school between science and arts) are also included as predictor variables.

The research was conducted at a prestigious university in Beijing, involving 238 undergraduate students enrolled in *Business English Writing* course. The course aimed to enhance students' academic writing skills in business studies by using LLMs for tasks such as register analysis, lexico-grammatical analysis, paraphrasing, text evaluation, generating written corrective feedback, as well as data analysis and reporting. A survey was administered to assess learners' motivation, performance expectancy, effort expectancy, social influence, facilitating condition, behavioral intention, and behavior use concerning LLMs adoption, and the data were analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM).

Based on the UTAUT theory, this study proposes an analytical framework for examining the factors influencing Chinese EFL learners' acceptance of LLMs in business academic writing. This theoretical framework is applicable for a comprehensive and objective investigation of the impact of new technologies on student agency in technology-enhanced language learning courses and can serve as a reference for subsequent researchers conducting further analyses in related fields. Additionally, unlike business English courses in other Asian countries, business English courses in Chinese universities tend to be larger in size, with class sizes typically ranging from 40 to 50 students (Wang and Xu, 2023). Moreover, within the same classroom, students come from both underdeveloped regions (such as county-level cities or rural areas) and developed regions (such as municipalities or prefecture-level cities). The significant wealth gap between students' families may lead to considerable differences in their motivations for learning English and their acceptance of new technologies. Furthermore, their prior exposure to computer courses during high school, along with the distinction between arts and science streams, may further amplify these differences. Therefore, this study can provide empirical insights for large EFL writing courses with significant background diversity. This study is structured into six sections: Section 1 introduces the research background, significance, and both theoretical and practical implications; Section 2 reviews the literature on the UTAUT model, motivation, and LLMs; Section 3 presents the hypotheses of the study;

Section 4 describes the data and research methods; Section 5 analyzes and discusses the statistical results; and Section 6 concludes the study.

2 Literature review

2.1 The UTAUT model and its application in technology-enhanced writing courses

The UTAUT model is a comprehensive framework that integrates elements from the Technology Acceptance Model (TAM), developed by Fred Davis (1989), and seven other models: the Theory of Reasoned Action, the Motivational Model, the Theory of Planned Behavior, a combined Theory of Planned Behavior/Technology Acceptance Model, the Model of Personal Computer Use, the Diffusion of Innovations Theory, and Social Cognitive Theory (Venkatesh et al., 2003). The UTAUT model aims to predict user acceptance of technology by considering various factors and individual differences, including four core variables: performance expectation, effort expectation, social influence, and facilitating conditions, along with moderating variables that affect these core variables. The constructs of the UTAUT model offer valuable insights into how these factors influence technology adoption in educational settings, particularly in the technology-enhanced writing courses.

One significant application of the UTAUT model in writing courses is its role in predicting students' acceptance of digital writing tools. Research indicates that performance expectancy significantly influences students' intentions to use writing software (Zeng, 2019). For instance, a study by Grassini et al. (2024) revealed that students who believed using technology would enhance their writing skills were more likely to engage with generative AI tools. This suggests that educators can foster technology acceptance by highlighting the benefits of writing tools in improving students' academic performance. Effort expectancy, which refers to the perceived ease of use, is another crucial factor in the adoption of technology in writing courses (Guggemos et al., 2020). Studies have shown that when students find digital tools intuitive and user-friendly, their willingness to utilize these resources increases (Budhathoki et al., 2024). This aligns with findings by Ferdousi (2022), who noted that simplifying the interface of writing software led to higher student engagement and satisfaction. Therefore, it is essential for educators to select and implement tools that minimize complexity to encourage their use in writing instruction. Social influence, which reflects the impact of peers and instructors on technology adoption, also plays a vital role in writing courses (Guggemos et al., 2020). This suggests that educators' attitudes towards technology can significantly affect students' perceptions and willingness to engage with digital writing resources. By actively promoting the integration of technology in writing courses, educators can create a supportive environment that fosters student engagement. Facilitating conditions, encompassing the resources and support available for technology use, are critical in the UTAUT framework as well (Hsu, 2023). Studies have shown that adequate access to technological resources, such as computers and internet connectivity, directly affects students' ability to engage with digital writing tools (Menon and Shilpa, 2023). This support is particularly vital in diverse classrooms where students may have varying levels of technological proficiency.

Moreover, the UTAUT model's emphasis on behavioral intention and use behavior allows for a nuanced understanding of how students interact with technology in writing contexts. Lin and Lai (2019)

suggested that the limited ability of behavioral intention to predict actual use of educational technology might be due to the moderating influence of users' self-regulation, with motivation being a key element of this self-regulation (Baumeister et al., 2007). This relationship underscores the importance of addressing motivational factors in technology adoption, as motivated students are more likely to integrate digital tools into their writing processes.

In conclusion, the application of the UTAUT model in technology-enhanced writing courses provides valuable insights into the factors influencing technology acceptance. By focusing on constructs such as performance expectancy, effort expectancy, social influence, and facilitating conditions, educators can better understand and facilitate students' engagement with digital writing tools. However, Menon and Shilpa (2023) suggested that the predictive power of the UTAUT variables is influenced by context. Thus, exploring its application in the context of LLMs is important. This study utilizes the UTAUT model to gain a better understanding of EFL learners' acceptance of LLMs for learning business-related English academic writing.

2.2 Motivation as a predictor and moderator in the UTAUT model

Educational disciplines have highlighted the critical role of learners' motivation, as it directly affects their academic performance, facilitates the transfer of acquired knowledge, and reinforces their persistence in learning (Stroet et al., 2015). Moreover, motivation plays a key role in language acquisition, being closely associated with learners' attitudes toward language learning and significantly influencing their efforts (Gardner, 2006; Wu, 2022). Research also indicates that motivation directly impacts users' intentions to adopt technology in educational settings. For example, Davis et al. (1992) suggested that intrinsic motivation significantly shapes users' perceptions of usefulness and ease of use—core components of the UTAUT framework. This finding is supported by Hsu (2023), who found that students with higher intrinsic motivation were more likely to perceive educational technologies as beneficial, thereby increasing their behavioral intention to use them.

Rafiee and Abbasian-Naghneh (2021) emphasized that learners' motivation should be included as a key factor in any research model within the field of education. Additionally, motivation has been recognized as a critical variable in explaining individuals' use of information technology (Roca and Gagné, 2008; Liu et al., 2024). Beyond serving as a predictor, motivation also acts as a moderator within the UTAUT model. For instance, Hsu (2023) found that motivation strengthens the relationship between social influence and EFL learners' intention to use LMOOCs. When learners are motivated, they are more responsive to social recommendations about technology use, suggesting that motivation amplifies the effects of the social influence construct in the UTAUT framework.

Based on the above research, understanding the role of motivation enhances insights into user intention and behavior regarding technology adoption in educational settings. Therefore, we incorporate motivation as both a predictor and a moderating variable within the UTAUT model. This approach measures the impact of motivation on EFL learners' intention and behavior in using LLMs for learning business-related English academic writing, as well as its moderating effects within the model.

2.3 LLMs and their application in teaching and learning academic writing

The advent of LLMs, such as OpenAI's GPT series and ERNIE Bot, has transformed the educational landscape, particularly in the teaching and learning of academic writing. These models harness vast amounts of text data to generate coherent and contextually relevant content, thereby providing valuable tools for students and educators alike. One of the most significant advantages of using LLMs in teaching and learning academic writing is their ability to offer immediate feedback. Studies have shown that real-time feedback can enhance students' academic writing skills by providing insights into the content, organization, vocabulary, grammar, syntax, cohesion, and mechanics of their writing (Xu and Wang, 2024a). For instance, LLMs can complement teacher assessment, which fosters a more iterative writing process (Lu et al., 2024). This interactive approach encourages students to engage with their writing actively, ultimately leading to better outcomes (Su et al., 2023).

Moreover, LLMs can serve as personalized writing assistants, catering to individual learning needs. Research indicates that ChatGPT is a valuable tool for learners at various proficiency levels to receive effective internal feedback (Tam, 2024). By adapting to the user's writing style and providing specific suggestions, LLMs can help students develop their unique voices while adhering to academic conventions (Xu and Wang, 2024b). This personalization is particularly beneficial in diverse classrooms where students may have varying levels of experience and expertise in writing. In addition, LLMs can enhance the writing process by assisting students in generating ideas and structuring their arguments. By querying an LLM, students can receive suggestions for relevant topics, thesis statements, arguments, counterarguments, and even outlines for their papers. This capability is especially useful for novice writers who may struggle with the initial stages of the academic writing process (Su et al., 2023). By streamlining the brainstorming phase, LLMs can reduce writing anxiety, allowing students to focus more on content development.

However, the use of LLMs in teaching and learning academic writing is not without challenges. Concerns have been raised about the potential for academic dishonesty (Rudolph et al., 2023; van Dis et al., 2023), particularly as students might over-rely on AI-generated content, thereby undermining their critical and creative thinking during the writing process (Barrot, 2023). To mitigate this risk, educators must emphasize the importance of critical thinking and originality in writing. Integrating LLMs into the curriculum should involve discussions about ethical use and the role of these tools as aids rather than crutches. Furthermore, while LLMs are powerful, they are not infallible. Issues such as biases in training data and the hallucination problems necessitate careful consideration and oversight in their application (OpenAI, 2022; Thorp, 2023; Barrot, 2023; Xu and Wang, 2024b). Educators must remain vigilant and guide students in critically evaluating the outputs of LLMs, fostering a mindset of discernment in the use of technological tools.

In conclusion, the application of LLMs in teaching and learning academic writing presents both opportunities and challenges. While previous studies have employed the TAM model to explore EFL learners' acceptance of LLMs in informal digital learning of English (Liu and Ma, 2024), to the best of our knowledge, there is limited research that has utilized the UTAUT model to analyze EFL learners'

acceptance of LLMs in learning business-related English academic writing, as well as the direct and moderating effects of motivation within the UTAUT model.

3 Hypotheses

3.1 Performance expectancy

Performance expectancy is the extent to which an individual believes that using a specific technology enhances their ability to perform tasks or achieve objectives (Venkatesh et al., 2003). It influences a person's intention to adopt that technology (Kumar and Bervell, 2019). In academic settings, performance expectancy plays a crucial role in determining the adoption of a particular technology (Budhathoki et al., 2024). In this research, performance expectancy refers to the belief among Chinese university students that using LLMs will improve their academic writing skills in business studies. Based on this, we hypothesize:

RH1: EFL learners' performance expectancy towards using LLMs would significantly affect their behavioral intention to use it.

3.2 Effort expectancy

Effort expectancy refers to the extent to which an individual perceives a particular technology as easy to use and requiring minimal effort (Venkatesh et al., 2003). It influences a person's intention to adopt that technology (Guggemos et al., 2020). In this study, effort expectancy represents the belief among students in Chinese higher education that LLMs are easy to use and require little effort. Based on this, we hypothesize:

RH2: EFL learners' effort expectancy towards using LLMs would significantly affect their behavioral intention to use it.

3.3 Social influence

Social influence refers to the extent to which an individual feels that important people in their life believe they should use a particular technology (Venkatesh et al., 2003). In this study, social influence pertains to how much students in Chinese universities believe that their peers, teachers, and other influential figures in their social network encourage them to use LLMs. Based on this, we hypothesize:

RH3: EFL learners' social influence towards using LLMs would significantly affect their behavioral intention to use it.

3.4 Facilitating conditions

Facilitating conditions refer to the extent to which individuals believe they have the necessary resources and support to use a particular technology effectively (Venkatesh et al., 2003). Studies have shown that facilitating conditions influence students' intentions to adopt new technologies (Kumar and Bervell, 2019; Hsu, 2023). In this

study, facilitating conditions refer to how much students believe they have sufficient access to LLMs, along with the knowledge and resources required to use them effectively in their learning. Based on this, we hypothesize:

RH4: EFL learners' facilitating conditions towards using LLMs significantly affect their use behavior.

3.5 Behavioral intention

Previous research indicates that behavioral intention has a positive, direct, and significant effect on technology usage behavior (Šumak and Šorgo, 2016; Budhathoki et al., 2024). This study suggests that performance expectancy, effort expectancy, and social influence directly and positively influence behavioral intention, which in turn positively affects usage behavior. Based on this, we hypothesize:

RH5: EFL learners' behavioral intention towards using LLMs would significantly affect their use behavior.

3.6 Motivation

Rafiee and Abbasian-Naghneh (2021) highlighted the importance of including learners' motivation as a key factor in any research model in the field of education. Moreover, motivation has been identified as a crucial factor in understanding individuals' use of information technology (Roca and Gagné, 2008; Liu et al., 2024). In addition to being a predictor, motivation also functions as a moderator within the UTAUT model (Hsu, 2023). Based on this, we hypothesize:

RH6: EFL learners' motivation significantly affects their behavioral intention towards using LLMs.

RH7: EFL learners' motivation significantly affects their LLMs use behavior.

RH8: EFL learners' motivation has a significant moderating effect on the relationship between the variables of UTAUT in LLMs settings.

3.7 Other factors

Due to the differences between business English courses in Chinese higher education and those in other countries and regions—particularly in terms of large class sizes and significant variations in both students' family economic backgrounds and previous learning experiences in high school (Wang and Xu, 2023)—we expanded the UTAUT model by including variables such as students' region (whether they come from developed areas like municipalities or prefecture-level cities, or less developed areas like county-level cities or rural regions), as well as their previous learning experiences (such as whether they have taken computer courses and their high school focus on science or arts), as predictor variables. Based on this, we hypothesize:

RH9: EFL learners' regional background significantly affects either their behavioral intention or use behavior regarding LLMs.

RH10: EFL learners' computer experience significantly affects either their behavioral intention or use behavior regarding LLMs.

RH11: EFL learners' high school stream significantly affects either their behavioral intention or use behavior regarding LLMs.

4 Data and methods

This study recruited 238 first-year undergraduate students from the School of Business at a prestigious university in Beijing to participate in a survey regarding their motivation and acceptance of LLMs in the *Business English Writing* course. The course aims to enhance EFL learners' English academic writing skills in the field of business through nine dimensions of assessment (vocabulary, grammar, orthographical control, genre format, cohesion & coherence, strategic competence, sociolinguistic competence, intercultural competence, and business knowledge) proposed by Wang and Fan (2020). In the course, we introduced participants to various LLMs to demonstrate their application methods and effects in business-related English academic writing, asking students to utilize them in their academic writing practice. Table 1 provides information about the writing module, teaching objectives, teaching methods, and the LLMs used.

In this study, we used SurveyMonkey to administer a questionnaire to 242 students enrolled in the *Business English Writing* course to investigate their motivation and acceptance of LLMs for learning English academic writing in the field of business, as well as their regional backgrounds and prior learning experiences. In the questionnaire, we explained the purpose of the study to the students, informing them that participation was completely voluntary and that they could withdraw at any time. Not participating or withdrawing would not affect their course grades. Ultimately, four students opted out of the survey, resulting in 238 valid responses. Throughout the research process, participants' identities were kept confidential.

This study used the key motivational factors proposed by Taguchi et al. (2009) to measure EFL learners' motivation toward LLMs in learning business-related English academic writing. These factors include integrativeness, instrumentality, attitudes, and two criterion measures, namely language choice preference and the learners' intended learning effort. Based on previous research (Venkatesh et al., 2003; Hsu, 2023), we incorporated motivation as a key latent variable into the UTAUT model and proposed the theoretical model for this study, as shown in Figure 1. This model proposes that behavioral intention serves as a crucial mediator between constructs such as motivation, performance expectancy, effort expectancy, social influence, and use behavior. Additionally, motivation, facilitating conditions, regional background, computer learning experience, and high school stream directly influence the use behavior, indicating that both psychological factors and external conditions play a key role in technology adoption. PLS-SEM was applied as the statistical technique to explore the structural relationships among variables. Furthermore, motivation was used as a moderating variable in the relationships between performance expectancy and behavioral intention, effort expectancy and behavioral intention, social influence

and behavioral intention, as well as facilitating conditions and use behavior.

4.1 Participants

We conducted this survey in the *Business English Writing* class at a prestigious university in Beijing, involving students from four business majors: finance, economics, business administration, and international business. A total of about 238 students participated in the experiment. Of the participants, roughly 60% came from developed regions of China, while around 40% were from underdeveloped regions. Additionally, around 16% had taken computer courses in high school, while the remaining 84% had not. Regarding their high school streams, about 25% were in the art stream, around 59% were in the science stream, and approximately 16% did not differentiate between the arts and science streams. All students participated voluntarily and were assured that the survey results would be used solely for academic research. All students participated voluntarily and were assured that the survey results would be used solely for academic research.

4.2 Measures

This study applied the L2 Motivational Self-System proposed by Taguchi et al. (2009) and the UTAUT model developed by Venkatesh et al. (2003) to measure Chinese EFL learners' motivation and acceptance of using LLMs in English academic writing for business studies. To better fit the scope of this research, we revised the motivation-related question items based on the English Learning Motivation Questionnaire (ELMQ) developed by Taguchi et al. (2009). Additionally, the UTAUT items were developed following studies by Venkatesh et al. (2003) and Hsu (2023). Specifically, 9 items assessed participants' motivation (e.g., 'Even if it's not required, I'm willing to use LLMs to learn business-related English academic writing.'). In the UTAUT model, five items evaluated participants' performance expectancy of LLMs (e.g., 'I find LLMs to be useful for learning business-related English academic writing.'). and three items examined their effort expectancy (e.g., 'I find LLMs for learning business-related English academic writing flexible and easy to use.'). Three items measured participants' social influence regarding LLM use (e.g., 'I would use LLMs for learning business-related English academic writing if my peers recommended it to me.'). and three items assessed the facilitating conditions (e.g., 'There is adequate training on the use of LLMs for learning business-related English academic writing in my university.'). Finally, participants' behavioral intention was measured with two items (e.g., 'I plan to use LLMs for learning business-related English academic writing.'). and behavior use was assessed with two items (e.g., 'I would enjoy using LLMs for learning business-related English academic writing.').

The questionnaire utilized a 6-point Likert scale, which included the categories: strongly agree, agree, slightly agree, slightly disagree, disagree, and strongly disagree. The composite reliability for each construct was above 0.7, confirming the reliability of the items (Fornell and Larcker, 1981a). To ensure that students understood the questionnaire accurately and to guarantee its validity, we first

TABLE 1 LLMs-assisted instruction in the *Business English Writing* course.

Module	Teaching objectives	Teaching method	LLMs ¹
Analysis of texts on the same business topic across different language registers	Develop students' register awareness in English academic writing within the field of business, and enhance their understanding of the differences in word and sentence styles between formal and informal English.	Use LLMs to generate two essays on the same business topic, one in a formal style and one in an informal style. Arrange group discussions for students to summarize the similarities and differences between the two essays. Additionally, compare these essays with their own writings to reflect on issues related to the language registers of their essays.	ERNIE Bot; ChatGLM; Kimi
Analysis of words, phrases and expressions related to specific business topics and case report writing	1. Enable students to master keywords and phrases related to specific business topics, as well as commonly used expressions related to case report writing. 2. Help students express their viewpoints more clearly and logically in English academic writing within the field of business.	Based on the writing topic and type, design different prompts to guide LLMs in generating commonly used words, phrases, and expressions related to specific business topics and case report writing.	ERNIE Bot; ChatGLM; Kimi
Paraphrase	Encourage students to use newly acquired English vocabulary, phrases, and common expressions to paraphrase existing simple paragraphs. By expanding and deepening the existing information in the paraphrase exercise, help students better master English academic writing in the field of business.	Design different prompts based on teaching tasks to make LLMs generate commonly used vocabulary related to relevant business topics, as well as original and paraphrased paragraphs for practice.	ERNIE Bot; ChatGLM; Kimi
Evaluating text & providing written corrective feedback	1. Guide students in using LLMs for self-assessment and to receive detailed written corrective feedback. 2. Help students more precisely identify areas for improvement and reflect on their writing.	By guiding students in designing different prompts, encourage them to use LLMs to evaluate their writing and obtain written corrective feedback on aspects such as vocabulary, grammar, spelling, structure, coherence, and originality.	ERNIE Bot; ChatGLM; Kimi
Data Analysis and Reporting	Guide students in using LLMs to analyze relevant data and present statistical results while writing case reports.	By guiding students in designing different prompts, encourage them to use LLMs to extract data from text, reform data, classify and score text, extract sentiment, etc.	ERNIE Bot; ChatGLM; Kimi
Wrap-up	By systematically reviewing the above teaching modules, students will gain a better understanding of the content they have learned.	By combining the activities of writing exercises, peer feedback, and teacher feedback, enhance students' ability to use LLMs in business English academic writing.	ERNIE Bot; ChatGLM; Kimi

¹We have incorporated three LLMs—ERNIE Bot, ChatGLM, and Kimi—into classroom teaching. Developed by leading Chinese high-tech companies, these LLMs represent the cutting edge of LLM development in China.

translated the questionnaire into Chinese and then back-translated it into English. Finally, we distributed the Chinese version of the questionnaire to participants via SurveyMonkey.

4.3 Data analysis

This study used Partial Least Squares Structural Equation Modeling (PLS-SEM) to conduct structural equation analysis. According to Hair et al. (2021), the advantages of PLS-SEM are as follows: (1) PLS-SEM can handle small sample sizes and non-normal data distributions effectively, making it suitable for exploratory research where large samples may not be feasible; (2) PLS-SEM emphasizes maximizing explained variance in the dependent constructs, which is beneficial for predictive modeling and understanding relationships in complex models; (3) PLS-SEM can effectively model both reflective and formative constructs, allowing for a more nuanced understanding of how different types of variables interact; (4) PLS-SEM is less sensitive to multicollinearity issues compared to traditional covariance-based SEM methods, making it a good choice when predictors are highly correlated; and (5) The output from PLS-SEM is relatively

straightforward, enabling researchers to easily interpret path coefficients and their significance, which aids in communicating findings.

Additionally, PLS-SEM is particularly useful for examining developing theories due to its higher statistical power (Hair et al., 2019). Therefore, this study adopted SmartPLS 4.0 for PLS-SEM analysis. The specific analysis steps were as follows: First, we assessed the overall reliability of the proposed model using R^2 and evaluated the reliability of the constructs using composite reliability (CR). Next, we employed the PLS-SEM algorithm in the software to estimate the path coefficients between variables and the moderating effect of motivation. Finally, we used bootstrapping to evaluate the confidence intervals and significance levels of the path coefficients, determining whether the path coefficients in the model were significant and providing the relevant p -values.

To assess the potential presence of common method bias (CMB) in our data, we employed Harman's single-factor technique (HSF), a widely used method for detecting CMB issues. The results of the exploratory factor analysis (EFA) showed that the first factor accounted for 38.53% of the total variance in the dataset. According to the guidelines in the literature, a "total variance explained" by the first component (or factor) extracted below 50% typically suggests that

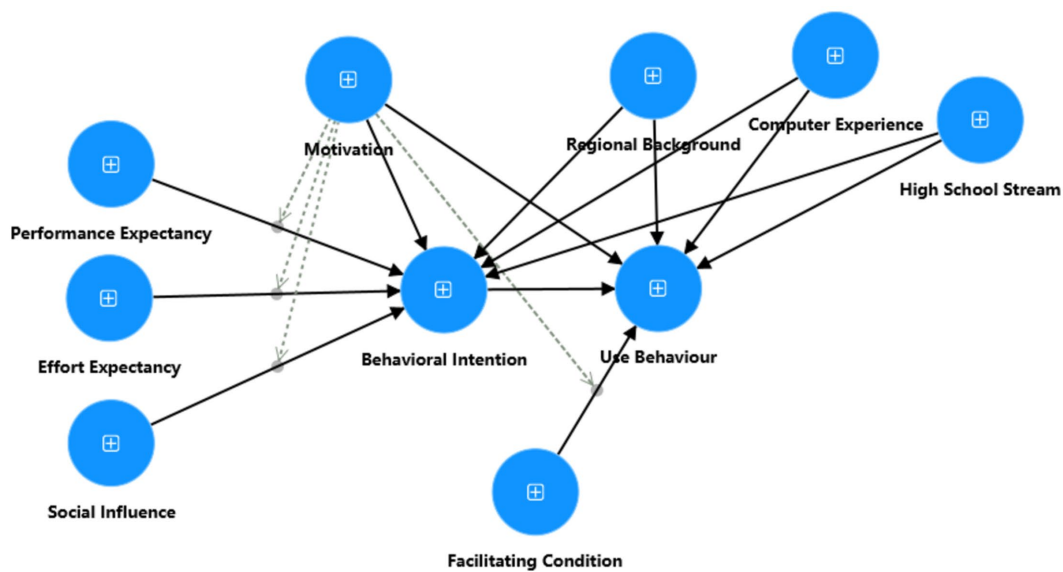


FIGURE 1
The proposed research model.

common method bias is not a serious concern (Podsakoff et al., 2003; Kock, 2020). Therefore, based on this result, we conclude that the dataset does not exhibit significant common method bias.

5 Results and discussion

Before analyzing the PLS-SEM results, it is crucial to verify the reliability and validity of the proposed research model. Table 2 provides the model and construct fit measures for the proposed model, focusing on R^2 and SRMR values to evaluate the model's predictive power, as well as Cronbach's α , Composite Reliability (ρ_c), and AVE to assess the reliability of the constructs. As shown in Table 2, The R^2 values for behavior intention (0.55) and use behavior (0.55) are relatively high, meaning that the UTAUT model explains a substantial portion of the variance in these constructs (Hair et al., 2011). The Standardized Root Mean Residual (SRMR) value of the proposed model is 0.07, which indicates that the model's predicted relationships closely match the observed data (Schermelleh-Engel et al., 2003). The R^2 and SRMR values demonstrate that the UTAUT model provides a strong explanation for participants' intentions to use and actual usage of LLMs in learning English academic writing. This suggests that key factors such as performance expectancy, effort expectancy, social influence, and motivation strongly predict both the likelihood of participants forming intentions to use LLMs and their actual usage.

To assess convergent validity, Cronbach's α , construct reliability (CR), and average variance extracted (AVE) were computed. A generally accepted rule is that a Cronbach's α of 0.7–0.8 indicates an acceptable level of reliability, while 0.8–0.9 indicates an ideal level (Tavakol and Dennick, 2011). All of our constructs meet this criterion, which means that they exhibit good internal consistency. Furthermore, previous studies suggest that a CR of 0.6 or higher (Fornell and Larcker, 1981b) and an AVE of 0.5 or higher (Hair et al., 2010) are

considered acceptable. As presented in Table 2, the AVEs for the seven latent constructs ranged from 0.50 to 0.84, meeting or exceeding the threshold, and the CR values for all constructs ranged from 0.70 to 0.91, exceeding the recommended value. These results confirm convergent validity and indicate good internal consistency for the constructs of the proposed model.

In addition, we calculate the Heterotrait-Monotrait ratio (HTMT) for motivation and UTAUT factors to assess discriminant validity between similar and different indicators. To establish discriminant validity, HTMT values must be lower than the threshold value of 0.90 (Gold et al., 2001; Teo et al., 2008). As shown in Table 3, all values are below 0.86. Furthermore, we calculated HTMT inference using bootstrapping with 5,000 subsamples. If the resulting interval is below 1, discriminant validity is confirmed (Henseler et al., 2015; Hernández-Perlines and Mancebo-Lozano, 2016). Our results also meet this criterion.

After confirming the fit of the proposed model, structural modeling was used to examine the relationships among the variables. As shown in Table 4, the findings of PLS-SEM reported that performance expectancy ($\beta_{\text{Performance expectancy}} = 0.29, p = 0.002$) and social influence ($\beta_{\text{Social Influence}} = 0.22, p = 0.003$) significantly influence EFL learners' behavioral intention to use LLMs in business academic writing. However, effort expectancy ($\beta_{\text{Effort Expectancy}} = 0.03, p = 0.588$) was not statistically significant. Moreover, participants' behavioral intention ($\beta_{\text{behavioral intention}} = 0.48, p = 0.000$) significantly influences their use behavior, while facilitating conditions ($\beta_{\text{facilitating condition}} = -0.06, p = 0.388$), regional background ($\beta_{\text{regional background}} = 0.12, p = 0.153$), computer experience ($\beta_{\text{computer experience}} = -0.05, p = 0.692$), and high school stream ($\beta_{\text{high school stream}} = -0.12, p = 0.417$) were not significantly associated with use behavior.

Our findings are partially in line with the studies by Guggemos et al. (2020). In their study, performance expectancy, effort expectancy, and social influence were strong predictors of users' intention to use digital tools in academic writing classes. However, this study found

that performance expectancy and social influence remain significant predictors of EFL learners' intention to use LLMs in academic writing courses, while the variable of effort expectancy was not. The possible explanation for this discrepancy may be related to the career planning and development of undergraduate students in Chinese universities. For most undergraduate students majoring in business in China, obtaining a master's degree or higher is required to secure better career development opportunities. Moreover, English academic writing ability is crucial for Chinese business undergraduates to qualify for a recommendation-based graduate admission or to

perform well in the graduate entrance exams (Wang and Xu, 2023). Therefore, as long as students anticipate that LLMs can help improve their English writing skills, they are willing to use this generative AI tool in the learning process, regardless of the effort required.

In this study, performance expectancy significantly influences users' behavioral intentions. This aligns with prior research that investigated students' intention to adopt generative AI tools using the UTAUT model (Bin-Nashwan et al., 2023; Budhathoki et al., 2024). The result indicates that Chinese EFL learners are more likely to use LLMs if they believe these tools will improve their academic writing performance. Therefore, in the context of higher education in China, educators should emphasize the practical benefits of LLMs in enhancing writing quality and efficiency to foster adoption.

Social influence also has a notable impact, as confirmed by several studies (Maican et al., 2019; Guggemos et al., 2020; Budhathoki et al., 2024). The findings suggest that recommendations from peers and endorsements from instructors influence Chinese EFL learners' intentions to use LLMs. Educators can leverage this by fostering a positive social atmosphere around LLM usage, encouraging peer discussions and collaborative work that support the adoption of these tools.

However, effort expectancy does not significantly affect behavioral intention in this study. This finding contradicts previous research (Guggemos et al., 2020; Kwak et al., 2022), and the possible reason, as mentioned above, may be related to the current career planning and development of undergraduate students in China. This study implies that learners' perceptions of ease of use are not as critical in this context. Learners appear to be more focused on the anticipated performance gains rather than the effort required to use LLMs. As a

TABLE 2 Model and construct fit.

Fit measures	Endogenous construct		
R ²	BI = 0.55; UB = 0.55		
SRMR	0.07		
	Construct reliability and validity		
Construct	Cronbach α	CR (rho_c)	AVE
Motivation	0.87	0.90	0.51
Performance Expectancy	0.87	0.91	0.67
Effort Expectancy	0.81	0.88	0.72
Social Influence	0.81	0.88	0.66
Facilitating Condition	0.80	0.70	0.50
Behavioral Intention	0.71	0.87	0.77
Use Behavior	0.82	0.91	0.84

TABLE 3 Heterotrait---Monotrait ratio (HTMT).

	Original sample (O)	Sample mean (M)	2.5%	97.5%
Motivation $\Rightarrow$ Behavioral Intention	0.79	0.796	0.58	0.94
Motivation $\Rightarrow$ Effort Expectancy	0.47	0.483	0.29	0.66
Motivation $\Rightarrow$ Facilitating Condition	0.13	0.182	0.10	0.32
Performance Expectancy $\Rightarrow$ Behavioral Intention	0.85	0.856	0.73	0.96
Performance Expectancy $\Rightarrow$ Effort Expectancy	0.77	0.771	0.65	0.86
Performance Expectancy $\Rightarrow$ Facilitating Condition	0.11	0.167	0.08	0.31
Performance Expectancy $\Rightarrow$ Motivation	0.70	0.709	0.51	0.84
Effort Expectancy $\Rightarrow$ Behavioral Intention	0.63	0.630	0.46	0.77
Social Influence $\Rightarrow$ Behavioral Intention	0.83	0.830	0.70	0.93
Social Influence $\Rightarrow$ Effort Expectancy	0.70	0.702	0.54	0.82
Social Influence $\Rightarrow$ Facilitating Condition	0.15	0.191	0.07	0.40
Social Influence $\Rightarrow$ Motivation	0.60	0.612	0.41	0.77
Social Influence $\Rightarrow$ Performance Expectancy	0.80	0.799	0.68	0.88
Facilitating Condition $\Rightarrow$ Behavioral Intention	0.03	0.121	0.03	0.28
Facilitating Condition $\Rightarrow$ Effort Expectancy	0.21	0.236	0.09	0.42
Use Behaviour $\Rightarrow$ Behavioral Intention	0.86	0.867	0.76	0.95
Use Behaviour $\Rightarrow$ Effort Expectancy	0.59	0.594	0.39	0.76
Use Behaviour $\Rightarrow$ Facilitating Condition	0.08	0.137	0.05	0.26
Use Behaviour $\Rightarrow$ Motivation	0.74	0.742	0.56	0.87
Use Behaviour $\Rightarrow$ Performance Expectancy	0.75	0.748	0.63	0.84
Use Behaviour $\Rightarrow$ Social Influence	0.72	0.719	0.50	0.87

TABLE 4 Path coefficients of PLS-SEM.

UTAUT	β	Standard deviation	p
Performance Expectance→Behavior Intention	0.29***	0.09	0.002
Effort Expectancy→Behavior Intention	0.03	0.06	0.588
Social Influence→Behavior Intention	0.22***	0.07	0.003
Motivation→Behavior Intention	0.29***	0.09	0.001
Regional Background→Behavior Intention	0.12	0.08	0.153
Computer Experience→Behavior Intention	−0.05	0.14	0.692
High School Stream→Behavior Intention	−0.12	0.15	0.417
Facilitating Condition→Use behavior	−0.06	0.07	0.388
Behavior Intention→Use Behavior	0.48***	0.09	0.000
Motivation→Use Behavior	0.33***	0.09	0.000
Regional Background→Use behavior	−0.09	0.08	0.266
Computer Experience→Use Behavior	0.04	0.10	0.645
High School Stream→Use Behavior	0.23	0.25	0.340

*** represent the 1% significance level.

result, Chinese educational institutions may not need to invest heavily in simplifying LLM interfaces or functionality, as perceived benefits and social influence are more influential factors.

Furthermore, behavioral intention is a strong predictor of actual LLM usage, aligning with previous studies that emphasize the importance of intention in technology adoption (Šumak and Šorgo, 2016; Budhathoki et al., 2024). This implies that once learners have a positive attitude toward using LLMs, they are highly likely to integrate them into their academic writing routines. Finally, factors such as facilitating conditions (e.g., external support systems), regional background, computer experience, and high school stream do not significantly influence usage, suggesting that once students form a positive intention, they are inclined to use LLMs regardless of other factors.

For the effect of participants' motivation as a predictor of their behavioral intention and use behavior toward LLMs in learning business-related English academic writing, as presented in Table 4, results showed that motivation was a significant factor influencing both behavioral intention ($\beta_{motivation1} = 0.29, p = 0.001$) and use behavior ($\beta_{motivation2} = 0.33, p = 0.000$). Our findings are consistent with those of Wang and Zhan (2020) and Strzelecki (2023), which indicate that Chinese EFL learners' motivation has a significant influence on both their intention to act and their actual behavior. The finding suggests that, beyond the traditional UTAUT constructs, motivation plays a crucial role in predicting how EFL learners engage with LLMs in learning business-related English academic writing. The inclusion of motivation aligns with extensions to the UTAUT model, emphasizing the importance of intrinsic factors like interest and desire in shaping EFL learners' technology acceptance. Therefore, educators should focus on enhancing motivational elements such as autonomy, competence, and relatedness to encourage students' adoption of LLMs in academic writing settings (Hsu, 2023).

This study also aims to explore the moderating effect of motivation on the variables of the UTAUT model. Table 5 shows the moderating effect of motivation on various variables within the UTAUT model,

TABLE 5 Moderating effect of motivation to variables of UTAUT.

UTAUT	β	Standard deviation	p
Motivation x Performance Expectancy → Behavior Intention	−0.01	0.07	0.858
Motivation x Effort Expectancy → Behavior Intention	0.03	0.06	0.616
Motivation x Social Influence → Behavior Intention	−0.04	0.06	0.499
Motivation x Facilitating Condition → Use Behavior	0.03	0.08	0.674

using β coefficients, standard deviations, and p -values to report the statistical significance of these relationships. The results indicated that motivation does not have a statistically significant influence on the relationships between the predictors and dependent variables under examination. Specifically, motivation does not significantly moderate the relationships between performance expectancy, effort expectancy, social influence, and behavioral intention. Similarly, for facilitating conditions, motivation does not significantly moderate its relationship with actual usage behavior.

The results suggest that motivation does not exhibit moderating effects on the UTAUT relationships, which may be because its influence on behavior could be less direct in the context of the model's constructs. According to Venkatesh et al. (2003), UTAUT primarily emphasizes the role of performance expectancy, effort expectancy, social influence, and facilitating conditions in shaping behavioral intention and usage behavior. Motivation, while important in influencing individual intentions and actions, might not significantly alter how these factors directly influence technology adoption. This finding was partially echoed in Hsu (2023), which suggests that individual differences, like intrinsic motivation, might not always have a strong moderating effect in the face of more dominant predictors.

Table 6 provides a summary of the examination of research hypotheses, detailing whether each hypothesis was supported or not. To sum up, performance expectancy, social influence, and motivation are significant predictors of EFL learners' behavioral intentions toward using LLMs in learning business-related English academic writing. Motivation and behavioral intention significantly influence EFL learners' actual use of LLMs in learning academic writing. The effects of other variables are not significant.

Table 6 provides a summary of the examination of the research hypotheses, detailing whether each hypothesis was supported. To summarize, performance expectancy, social influence, and motivation are significant predictors of EFL learners' behavioral intentions toward using LLMs in learning business-related English academic writing. Motivation and behavioral intention significantly influence EFL learners' actual use of LLMs in academic writing. On the other hand, the effects of other variables are not significant.

6 Conclusion and contribution

This study investigated the factors influencing EFL learners' adoption of LLMs in learning business-related English academic writing by applying the UTAUT model and incorporating motivation,

TABLE 6 Summary of research hypotheses examination.

Research hypotheses	Results
RH1: EFL learners' performance expectancy toward using LLMs would significantly affect their behavioral intention to it.	Supported
RH2: EFL learners' effort expectancy toward using LLMs would significantly affect their behavioral intention to it.	Not Supported
RH3: EFL learners' social influence toward using LLMs would significantly affect their behavioral intention to it.	Supported
RH4: EFL learners' facilitating conditions using toward LLMs would significantly affect their use behavior.	Not Supported
RH5: EFL learners' behavioral intention toward using LLMs would significantly affect their use behavior.	Supported
RH6: EFL learners' motivation significantly affects their behavioral intention toward using LLMs.	Supported
RH7: EFL learners' motivation significantly affects their LLMs use behavior.	Supported
RH8: EFL learners' motivation has a significant moderating effect on the relationship between UTAUT variables in LLMs settings.	Not Supported
RH9: EFL learners' regional background significantly affects either their behavioral intention or use behavior regarding LLMs.	Not Supported
RH10: EFL learners' computer experience significantly affects either their behavioral intention or use behavior regarding LLMs.	Not Supported
RH11: EFL learners' high school stream significantly affects either their behavioral intention or use behavior regarding LLMs.	Not Supported

regional background, and previous learning experiences (computer experience and high school streams) as key variables. The research primarily aimed to understand the effects of performance expectancy, effort expectancy, social influence, motivation, regional background, and previous learning experiences on both behavioral intention and actual usage of LLMs among EFL learners in academic writing contexts. Additionally, the study sought to explore the moderating effect of motivation on these relationships.

The major findings indicate that performance expectancy and social influence significantly predict EFL learners' behavioral intention to use LLMs in their business-related English academic writing. Motivation also emerged as a significant predictor of both behavioral intention and use behavior, confirming its central role in encouraging the adoption of educational technologies in academic writing contexts. On the other hand, effort expectancy did not significantly influence behavioral intention, suggesting that learners might prioritize the perceived benefits of LLMs over the ease of using them. Interestingly, the facilitating conditions, which refer to the availability of resources and support for using LLMs, did not have a significant effect on learners' actual use behavior. This could imply that, in the context of this study, learners already have adequate access to the resources needed to use LLMs, or that their motivation and the perceived benefits of the technology outweigh the importance of external support in learning academic writing. The moderating role of motivation was not supported by the empirical evidence, as motivation did not significantly alter the relationships between key UTAUT variables and behavioral intention. Moreover, although there are significant differences among students in terms of regional background and previous

learning experience, these differences did not significantly impact their behavioral intention and use behavior.

This study contributes to the literature by providing both theoretical and practical implications for the use of cutting-edge technology in EFL writing courses. This study's theoretical implications, grounded in UTAUT theory, offer an analytical framework to examine the factors that influence Chinese EFL learners' acceptance of new technology in writing courses. This framework incorporates variables such as motivation, regional background, and differences in previous learning experience, which are common in EFL classrooms in Chinese higher education. It is highly applicable for a comprehensive and objective study of the agency of Chinese EFL learners in technology-assisted writing courses and can serve as a reference for researchers conducting further analysis in related fields.

Equally important, the findings also provide several practical implications for improving EFL learners' acceptance of LLMs in learning academic writing, particularly in the field of business. First, interventions aimed at increasing performance expectancy and social influence should be prioritized, as these are key factors driving learners' intention to use LLMs. Educators could emphasize the performance benefits of LLMs in enhancing writing skills or create a collaborative environment where peer encouragement fosters greater technology adoption. For instance, in the intervention, we asked students to work in groups to use LLMs to search for specific business terms (e.g., "digital transformation," "sustainability," "green finance," "trade dispute/war") and analyzed their frequency and usage patterns over a set period (e.g., the past five years). Next, students were required to create a timeline showing how the usage of each term has evolved, and reflect on why certain terms have gained or lost prominence based on business events or trends. Finally, students worked in groups to discuss and present the implications of these trends.

Through group collaboration, reporting, and reflection activities like these, students can gain insight into how their peers use LLMs to learn business English vocabulary and their reflections on its usage, thereby further enhancing their acceptance of this technology.

Second, motivation stands out as a crucial factor not only in shaping learners' intention but also in driving actual usage. Therefore, fostering intrinsic motivation through engaging, relevant, and supportive learning environments is essential. Specifically, to increase EFL learners' motivation toward using LLMs in business English academic writing, educators can adopt several effective strategies. First, goal-setting is essential; by providing clear, achievable objectives—such as improving writing quality or mastering key business terms—learners stay focused and motivated. Immediate feedback also plays a crucial role, allowing students to track their progress and adjust their approach accordingly, which reinforces their self-efficacy. In addition, educators can foster a supportive learning community by encouraging peer collaboration. Group activities, where students use LLMs to explore business-related vocabulary or analyze trends, can boost motivation through social interaction and shared learning experiences. Gamification elements, such as small rewards or recognition for achievements, can introduce an element of fun and competition, keeping learners engaged. Finally, real-world relevance is key. Educators should emphasize how LLMs can enhance learners' business writing skills, showing them the direct impact on future academic and professional success, making the learning process more meaningful.

Our analysis, however, has limitations. First, the study focused solely on EFL learners in the context of academic writing in the field

of business, which limits the generalizability of the findings to promoting the use of LLMs in academic writing in other fields. Second, although motivation was treated as a moderating variable, the results showed no support for its moderating effects, which may be due to the relatively small sample size. Finally, the study relied on self-reported data, which may be subject to biases such as social desirability or inaccurate self-assessment.

In the future, we plan to explore how these factors manifest across different academic fields and technologies to determine if the findings are consistent in varied educational settings. Additionally, we aim to increase the sample size and include more moderating variables, such as students' learning style and metacognitive awareness, to gain a more comprehensive and robust understanding of the moderating effects of motivation and other factors within the UTAUT model. Future research could also incorporate more objective measures, such as actual usage data or performance assessments, to provide a clearer picture of how motivation and UTAUT factors influence EFL learners' intentions and adoption of LLMs in academic writing contexts.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee of China University of Political Science and Law. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

QW: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

References

- Barrot, J. S. (2023). Using ChatGPT for second language writing: pitfalls and potentials. *Assess. Writ.* 57:100745. doi: 10.1016/j.asw.2023.100745
- Baumeister, R. F., Vohs, K. D., and Tice, D. M. (2007). The strength model of self-control. *Curr. Dir. Psychol. Sci.* 16, 351–355. doi: 10.1111/j.1467-8721.2007.00534.x
- Bin-Nashwan, S. A., Sadallah, M., and Bouteraa, M. (2023). Use of ChatGPT in academia: academic integrity hangs in the balance. *Technol. Soc.* 75:102370. doi: 10.1016/j.techsoc.2023.102370
- Budhathoki, T., Zirar, A., Njaya, E. T., and Timsina, A. (2024). ChatGPT adoption and anxiety: a cross-country analysis utilising the unified theory of acceptance and use of technology (UTAUT). *Stud. High. Educ.* 49, 831–846. doi: 10.1080/03075079.2024.2333937
- Davis, F. D. (1989). "Perceived usefulness, perceived ease of use, and user acceptance of information technology". *Management Information System Quarterly*. 13, 983–1003. doi: 10.2307/249008
- Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace. *J. Appl. Soc. Psychol.* 22, 1111–1132. doi: 10.1111/j.1559-1816.1992.tb00945.x
- Ferdousi, B. (2022). Acceptance of technology and academic writing: analyze in perspective of theoretical models. *J. Syst. Cybernet. Informatics* 20, 94–117. doi: 10.54808/JSCI.20.05.94
- Fornell, C., and Larcker, D. F. (1981a). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50. doi: 10.1177/002224378101800104
- Fornell, C., and Larcker, D. F. (1981b). Structural equation models with unobservable variables and measurement error: algebra and statistics. *J. Mark. Res.* 18, 382–388. doi: 10.1177/002224378101800313
- Gardner, R. C. (2006). The socio-educational model of second language acquisition: a research paradigm. *EUROSLA Yearb.* 6, 237–260. doi: 10.1075/eurosla.6.14gar

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was supported by Humanities and Social Sciences Youth Foundation of Ministry of Education of China (Grant No.: 24YJC740066).

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. This study aims to measure EFL learners' motivation and acceptance of using large language models (the most important product of generative AI) in English academic writing in business studies. Therefore, in the classroom instruction on which this study is based, we required participating students to use three large language models—ERNIE Bot, ChatGLM, and Kimi—to complete various academic writing tasks.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1514545/full#supplementary-material>

- Gold, A. H., Malhotra, A., and Segars, A. H. (2001). Knowledge management: an organizational capabilities perspective. *J. Manag. Inf. Syst.* 18, 185–214. doi: 10.1080/07421222.2001.11045669
- Grassini, S., Aasen, M. L., and Møgelvang, A. (2024). Understanding university students' acceptance of ChatGPT: insights from the UTAUT2 model. *Appl. Artif. Intell.* 38:2371168. doi: 10.1080/08839514.2024.2371168
- Guggemos, J., Seufert, S., and Sonderegger, S. (2020). Humanoid robots in higher education: evaluating the acceptance of pepper in the context of an academic writing course using the UTAUT. *Br. J. Educ. Technol.* 51, 1864–1883. doi: 10.1111/bjet.13006
- Hair, J. F., Celsi, D. M., Ortinau, J., and Bush, R. P. (2010). *Essentials of marketing research*. New York: McGraw-Hill.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., and Ray, S. (2021). Partial least squares structural equation modeling (PLS-SEM) using R: A workbook. Berlin: Springer Nature.
- Hair, J. F., Ringle, C. M., and Sarstedt, M. (2011). PLS-SEM: indeed a silver bullet. *J. Mark. Theory Pract.* 19, 139–152. doi: 10.2753/MTP1069-6679190202
- Hair, J. F., Risher, J. J., Sarstedt, M., and Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *Eur. Bus. Rev.* 31, 2–24. doi: 10.1108/EBR-11-2018-0203
- Henseler, J., Ringle, C. M., and Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* 43, 115–135. doi: 10.1007/s11747-014-0403-8
- Hernández-Perlines, F., and Mancebo-Lozano, E. (2016). Conditional mediation of competitive strategy and environment in international entrepreneurial orientation of family businesses. *European J. Family Bus.* 6, 86–98. doi: 10.1016/j.ejfb.2017.05.001
- Holmes, W., and Miao, F. (2023). *Guidance for generative AI in education and research*. Paris: UNESCO Publishing.
- Hsu, L. (2023). EFL learners' self-determination and acceptance of LMOOCs: the UTAUT model. *Comput. Assist. Lang. Learn.* 36, 1177–1205. doi: 10.1080/09588221.2021.1976210
- Kock, N. (2020). Harman's single factor test in PLS-SEM: checking for common method bias. *Data Anal. Persp. J.* 2, 1–6.
- Kumar, J., and Bervell, B. (2019). Google classroom for mobile learning in higher education: modelling the initial perceptions of students. *Educ. Inf. Technol.* 24, 1793–1817. doi: 10.1007/s10639-018-09858-z
- Kwak, Y., Seo, Y. H., and Ahn, J. W. (2022). Nursing students' intent to use AI-based healthcare technology: path analysis using the unified theory of acceptance and use of technology. *Nurse Educ. Today* 119:105541. doi: 10.1016/j.nedt.2022.105541
- Lin, J. W., and Lai, Y. C. (2019). User acceptance model of computer-based assessment: moderating effect and intention-behavior effect. *Australas. J. Educ. Technol.* 35:4684. doi: 10.14742/ajet.4684
- Liu, G., and Ma, C. (2024). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innov. Lang. Learn. Teach.* 18, 125–138. doi: 10.1080/17501229.2023.2240316
- Liu, G. L., Zhang, Y., and Zhang, R. (2024). Examining the relationships among motivation, informal digital learning of English, and foreign language enjoyment: an explanatory mixed-method study. *ReCALL* 36, 72–88. doi: 10.1017/S0958344023000204
- Lu, Q., Yao, Y., Xiao, L., Yuan, M., Wang, J., and Zhu, X. (2024). Can ChatGPT effectively complement teacher assessment of undergraduate students' academic writing? *Assessment and evaluation in higher education*. 2024, 616–633. doi: 10.1080/02602938.2024.2301722
- Maican, C. I., Cazan, A.-M., Lixandriou, R. C., and Dolveac, L. (2019). A study on academic staff personality and technology acceptance: the case of communication and collaboration applications. *Comput. Educ.* 128, 113–131. doi: 10.1016/j.compedu.2018.09.010
- Menon, D., and Shilpa, K. (2023). Chatting with ChatGPT: Analyzing the factors influencing users' intention to Use the Open AI's ChatGPT using the UTAUT model. *Heliyon*, 9. doi: 10.1016/j.heliyon.2023.e20962
- Murray, N., Liddicoat, A. J., Zhen, G., and Mosavian, P. (2023). Constraints on innovation in English language teaching in hinterland regions of China. *Language Teaching Research*, 27, 1246–1267. doi: 10.1177/1362168820979855
- OpenAI (2022). *Introducing ChatGPT*. OpenAI. Available online at: <https://openai.com/blog/chatgpt>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., and Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J. Appl. Psychol.* 88, 879–903. doi: 10.1037/0021-9010.88.5.879
- Rafiee, M., and Abbasian-Naghneh, S. (2021). E-learning: development of a model to assess the acceptance and readiness of technology among language learners. *Comput. Assist. Lang. Learn.* 34, 730–750. doi: 10.1080/09588221.2019.1640255
- Roca, J. C., and Gagné, M. (2008). Understanding e-learning continuance intention in the workplace: a self-determination theory perspective. *Comput. Hum. Behav.* 24, 1585–1604. doi: 10.1016/j.chb.2007.06.001
- Rudolph, J., Tan, S., and Tan, S. (2023). ChatGPT: bullshit spewer or the end of traditional assessments in higher education? *J. Appl. Learn. Teach.* 6, 342–363. doi: 10.37074/jalt.2023.6.1.9
- Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol. Res. Online* 8, 23–74. doi: 10.23668/psycharchives.12784
- Stroet, K., Opdenakker, M. C., and Minnaert, A. (2015). What motivates early adolescents for school? A longitudinal analysis of associations between observed teaching and motivation. *Contemp. Educ. Psychol.* 42, 129–140. doi: 10.1016/j.cedpsych.2015.06.002
- Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interact. Learn. Environ.*, 32, 5142–5155. doi: 10.1080/10494820.2023.2209881
- Su, Y., Lin, Y., and Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assess. Writ.* 57:100752. doi: 10.1016/j.asw.2023.100752
- Šumak, B., and Šorgo, A. (2016). The acceptance and use of interactive whiteboards among teachers: differences in UTAUT determinants between pre- and post-adopters. *Comput. Hum. Behav.* 64, 602–620. doi: 10.1016/j.chb.2016.07.037
- Taguchi, T., Magid, M., and Papi, M. (2009). “The L2 motivational self system among Japanese, Chinese, and Iranian learners of English: a comparative study” in *Motivation, language identity and the L2 self*, (Eds.), Z. Dörnyei and E. Ushioda. (Bristol: Multilingual Matters), 66–97. doi: 10.21832/9781847691293-005
- Tam, A. C. F. (2024). Interacting with ChatGPT for internal feedback and factors affecting feedback quality. *Assessment and evaluation in higher education*, 1–17. doi: 10.1080/02602938.2024.2374485
- Tavakol, M., and Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53. doi: 10.5116/ijme.4dfb.8dfd
- Teo, T. S. H., Srivastava, S. C., and Jiang, L. (2008). Trust and electronic government success: an empirical study. *J. Manag. Inf. Syst.* 25, 99–132. doi: 10.2753/MIS0742-1222250303
- Thorp, H. H. (2023). ChatGPT is fun but not an author. *Science* 379:313. doi: 10.1126/science.adg7879
- van Dis, E. A., Bollen, J., Zuidema, W., Van Rooij, R., and Bockting, C. L. (2023). ChatGPT: five priorities for research. *Nature* 614, 224–226. doi: 10.1038/d41586-023-00288-7
- Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *MIS Q.* 27, 425–478. doi: 10.2307/30036540
- Wang, L., and Fan, J. (2020). Assessing business English writing: the development and validation of a proficiency scale. *Assess. Writ.* 46:100490. doi: 10.1016/j.asw.2020.100490
- Wang, Q., and Xu, J. (2023). Interdisciplinary research on translation teaching and talent training in the new era: taking business translation teaching as an example. *Foreign Lang. China* 4, 89–97. doi: 10.13564/j.cnki.issn.1672-9382.2023.04.005
- Wang, W., and Zhan, J. (2020). The relationship between English language learner characteristics and online self-regulation: a structural equation modeling approach. *Sustain. For.* 12:3009. doi: 10.3390/su12073009
- Wu, X. (2022). Motivation in second language acquisition: a bibliometric analysis between 2000 and 2021. *Front. Psychol.* 13:1032316. doi: 10.3389/fpsyg.2022.1032316
- Wu, X. H., Fang, S., and Lai, I. K. W. (2019). “Undergraduate student's acceptance of a situational and interactive hotel English learning APP: an empirical study based on the extension of UTAUT” in *International conference on Technology in Education* (New York: Springer), 203–212.
- Xu, J., and Wang, Q. (2024a). Applying neural machine translation and ChatGPT in the teaching of business English writing. *Transl. Transl. Multilingual Contexts* 11, 1–39.
- Xu, J., and Wang, Q. (2024b). Innovation and exploration in foreign language teaching in the era of digital humanities: a case study of large language models in ESP writing instruction. *Foreign Lang. China* 6, 89–97.
- Zeng, J. (2019). A study on user acceptance of mobile-assisted language learning based on UTAUT2. *Technol. Enhanced Foreign Lang. Educ.* 6, 16–24.
- Zhang, H. (2017). Opportunity or new poverty trap: rural-urban education disparity and internal migration in China. *China Econ. Rev.* 44, 112–124. doi: 10.1016/j.chieco.2017.03.011



OPEN ACCESS

EDITED BY

Xinghua Liu,
Shanghai Jiao Tong University, China

REVIEWED BY

Yancheng Yang,
Hunan University, China
Veronika Makarova,
University of Saskatchewan, Canada
Chenghui Chen,
Nanjing University of Finance and Economics,
China

*CORRESPONDENCE

Chuanbo Ji
✉ jichuanbo@pku.edu.cn

[†]These authors share first authorship

RECEIVED 22 October 2024

ACCEPTED 24 February 2025

PUBLISHED 01 April 2025

CITATION

Gong X, Liu R and Ji C (2025) Citation practices in applied linguistics: a comparative study of Chinese expert and novice authors. *Front. Psychol.* 16:1515323. doi: 10.3389/fpsyg.2025.1515323

COPYRIGHT

© 2025 Gong, Liu and Ji. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Citation practices in applied linguistics: a comparative study of Chinese expert and novice authors

Xue Gong^{1†}, Ruoxi Liu^{2†} and Chuanbo Ji^{2*}

¹School of Humanities and Law, Hebei University of Technology, Tianjin, China, ²School of Chinese as a Second Language, Faculty of Humanities, Peking University, Beijing, China

Citation practices are crucial in academic discourse for both knowledge construction and interpersonal interaction. While prior research in academic English has explored citation practices among expert and novice authors, there is a notable gap in studies focusing on Chinese academic papers. Moreover, it remains uncertain whether insights from English-language corpora can be extrapolated to other linguistic contexts. This study presents a comparative analysis of citation practices among expert and novice authors within the field of Chinese Applied Linguistics. Utilizing a corpus of 715,000 Chinese words, we analyzed academic papers authored by both groups. Our findings reveal that citation practices between expert and novice authors are largely comparable. Specifically, integral citations were more prevalent than non-integral citations, with the cited authors predominantly occupying the subject position. In terms of citation form, the four types employed, in descending order of frequency, were summary, block quote, generalization, and quote. The analysis of reporting markers showed a predominance of discourse markers, followed by research markers, with cognitive markers being the least frequent. Notably, novice authors demonstrated certain deficiencies compared to their expert counterparts, including an overreliance on integral citations, a reduced use of generalization and block quote citations, and limited integration of information regarding reporting markers.

KEYWORDS

Chinese applied linguistics, citation practices, expert authors, novice authors, academic writing

1 Introduction

Citations are a fundamental component of academic discourse, facilitating both the dissemination of ideas and the exchange of research findings within scholarly communities. The functions of citations can be categorized into three primary aspects: knowledge construction, intertextuality, and interpersonal interaction. At the level of knowledge construction, the generation of new knowledge relies on the integration of a shared disciplinary framework. Incorporating previous research findings is essential for constructing new insights, positioning citations as a vital tool for presenting and advancing scientific knowledge (Hyland, 1999). From an intertextual standpoint, citations extend the discourse beyond the immediate text, allowing for the integration of the current study with prior research in the field. This positions the research within the broader scholarly landscape (Hyland and Jiang, 2019). In terms of interpersonal interaction, citations facilitate engagement with two key groups. First, within the academic community, citations are used to assess existing research—through agreement, critique, or neutrality. These interactions foster dialogic relationships to advance disciplinary knowledge (Thompson and Ye, 1991). Moreover, to persuade readers effectively,

citations of relevant studies are necessary to support the proposed arguments (Hyland, 1999; Hyland, 2010; Mansourizadeh and Ahmad, 2011; Chen and Zhang, 2017).

Citation practices have been a focal point across disciplines such as applied linguistics, sociology of knowledge, and information science (Swales, 1986; Bazerman, 1988; White, 2004). In applied linguistics, key areas of investigation include the forms and functions of citations (e.g., Harwood, 2009; Swales, 2014), interdisciplinary variations in citation practices (e.g., Hyland, 1999), differences between native and non-native English authors (e.g., Li and Zhang, 2021), and contrasts between expert and novice authors (e.g., Marti et al., 2019).

The majority of existing studies have focused on the examination of citations in academic English writing, with comparatively limited attention directed toward Chinese papers. Indeed, there are notable differences between citation practices in academic English and those in academic Chinese. For instance, while Arizavi and Choubasaz (2021) observed a greater prevalence of non-integral citations in academic English papers, Peng (2019) reported the opposite trend, namely that scholars who are trained in China exhibited a stronger tendency to use integral citations in their English papers. This also indicates that the influence of the mother tongue on citation practices is a factor that should not be overlooked. It is therefore necessary to analyse Chinese-language papers in order to gain new insights.

2 Literature review

2.1 Research on citation practice in academic writing

A substantial amount of research has been conducted on citation practices in academic writing. Some studies have examined the form and function of citation practices (Swales, 1986; Hyland, 1999; Thompson and Ye, 1991; Petrić, 2007; Mansourizadeh and Ahmad, 2011). Furthermore, studies have been conducted that have explored differences in citation practices in papers from different disciplines (Hyland, 1999; Hu and Wang, 2014; Wang and Hu, 2022), and differences among academic writers with different cultural backgrounds (Li, 2011; Cui and Cheng, 2014; Peng, 2019), and differences among writers with English as a second language versus native English speakers (Sun, 2009; Lou, 2011; Li, 2012; Shi, 2013; Lee et al., 2018; Li and Zhang, 2021), and differences in different genres (e.g., introductions, methodology, results, and discussion genres for empirical papers) (Martínez, 2008; Kwan and Chan, 2014; Zhang, 2022; Zhang, 2023).

In relation to the manner of citation practice, the extant research can be summarized as follows:

Embedding method: Swales (1986) classified citations as integral or non-integral, based on the position of the quoted person within or outside of the sentence. The academic community has endorsed this classification and it has been adopted by subsequent studies related to citation practices in academic writing (Hyland, 1999; Charles, 2006; Mansourizadeh and Ahmad, 2011; Samraj, 2013; Zhang, 2023; Mu, 2024). Furthermore, studies have been conducted that refined the categorization based on the syntactic position of the cited authors (Hyland, 1999; Thompson and Tribble, 2001). To illustrate this,

Hyland (1999) categorized integral citations into the cited author as subject, non-subject, and situated in a noun phrase.

Citation form: Hyland (1999) divides the citation form into four categories: summary, generalization, quote, and block quote. Summary and generalization are indirect citations. Summary means that the material quoted is attributed to one source. Generalization means that the material is attributed to two or more sources. Quote and block quote are direct citations. Quote is a short direct quotation (three or more words). Block quote refers to extensive use of the original wording, set out as indented blocks. Borg (2000) and Petrić (2012) further classify direct quotations into three categories: quotation fragments (stretches of textual borrowing shorter than a T-unit), short quotations (T-units shorter than 40 words), extended quotations (quotations longer than 40 words).

Reporting markers: Thompson and Ye (1991) classified reporting markers into three categories, namely research markers (e.g., observe), cognitive markers (e.g., believe), and discourse markers (e.g., discuss). Hyland (1999) and Liu et al. (2021) followed this categorization.

2.2 Variation of citation practice by writer expertise

The differences in citation practices among writing groups with varying levels of expertise can be observed in five main ways.

First, writing groups with higher levels of expertise tend to have a higher citation density than those with lower levels of expertise. For example, Lombardi (2021) study demonstrated that high-level writers cite more frequently than low-level writers.

Second, with regard to the embedding method and citation form, writing groups with higher levels of expertise tend to employ a greater number of non-integral citations (Mansourizadeh and Ahmad, 2011; Ahn and Oh, 2024), and they tend to introduce shorter segments of source material (Lombardi, 2021).

Third, the use of reporting markers is more diverse in terms of the reporting verbs employed by writers with higher levels of expertise (Lombardi, 2021).

Fourth, writers with higher levels of expertise tend to evaluate the cited content and express their personal stance in their citation practices (Wette, 2018; Zhang, 2023). For example, Lombardi (2021) study demonstrated that high-level writers are more likely to attach personal evaluations to reporting markers than their less experienced counterparts. In contrast, low-level writers tend to avoid evaluative citations (Li and Zhang, 2021). Furthermore, doctoral students have been observed to utilize evaluative citations more frequently than their master's counterparts (Zhao and Zhan, 2020).

Fifth, in terms of the function of citations, expert authors are more proficient in employing citations to serve their communicative purposes. For instance, doctoral students are more likely to cite sources than master's students (Li and Zhang, 2021). It has been demonstrated that experts are more proficient in substantiating their personal discourses through the use of citations (Mansourizadeh and Ahmad, 2011; Mu, 2024). Additionally, experts are more inclined to engage in comparative analysis of research findings through the utilization of citations within the discussion section (Samraj, 2013). Furthermore, studies have examined the utilization of citations in research grant applications by novice authors, revealing that novice

authors are in a transitional phase from student to researcher and have not yet developed the academic writing skills and competencies comparable to those of expert authors (Fazel and Shi, 2015).

A review of the existing research reveals three areas that warrant further investigation. First, the issue of expertise influencing citation practice is still in its infancy. The majority of articles were published subsequent to 2020, and the total number of articles is relatively limited. It is noteworthy that while existing academic writing textbooks address citation practices, they tend to focus on reminding writers to avoid plagiarism through proper citation, rather than on choosing the most appropriate form of citation for communicative purposes. This makes it challenging for novice writers to obtain effective guidance from these textbooks. Second, the extant studies utilize English papers as the corpus, with fewer studies focusing on citation practices in Chinese papers and an even smaller number of studies on the citation practices of novice Chinese academic writers. It has been demonstrated that there are differences between the academic citation practices of English and Chinese (Arizavi and Choubasaz, 2021; Peng, 2019). Consequently, it is necessary to re-examine the latter. Third, there are already established studies on the citation practices of novice authors, which compare dissertations with journal papers. For example, Li and Zhang (2021) and Ahn and Oh (2024) have already conducted such studies. However, it should be noted that dissertations and journal papers belong to two different genres. Therefore, further research is needed to explore whether citation practices can be compared across genres.

In light of the aforementioned background, this paper seeks to address two research questions by constructing a corpus of expert and novice academic Chinese journal papers:

- 1 Does the number of citations vary according to the level of expertise of the writers?
- 2 Does embedding method, citation form, and reporting markers of citations vary according to the level of expertise of the writer?

3 Methods

This study is based on a corpus of 190 journal papers, comprising a total of 715,000 words, drawn from two distinct writing groups with varying levels of expertise (experts/novices) in their respective fields. In consideration of the disciplinary variation in citation practices (Hyland, 1999), the corpus for this research is limited to that of applied linguistics. The rationale for selecting this discipline is based on the researchers' familiarity with it, which ensures more reliable findings.

3.1 Data collection: the corpora

The corpora for this study are categorized into two segments: expert-authored papers and novice-authored papers.

The corpus of expert authors' papers was created in the following manner. The citation analysis feature of China National Knowledge Infrastructure (CNKI) was employed to examine the most highly-cited authors in the field from five core journals between 2015 and 2020. The journals in question are *Chinese Teaching in the World*, *Language Teaching and Linguistic Studies*,

Applied Linguistics (Yuwan Wenzhi Yingyong), *Chinese Language Learning*, and *Chinese Linguistics*. The 10 most prolific authors from each of the aforementioned journals were selected as potential experts. From this initial list, authors were further filtered based on their substantial individual publication record and significant recognition within the academic community. Ultimately, a total of 95 papers authored by 14 expert authors were chosen to establish the expert papers' corpus, comprising 379,000 words.

The corpus of novice authors' papers was created in the following manner. The data selected for the novice authors' papers were sourced from the Graduate Forum organized by the School of Chinese as a Second Language at Peking University, spanning the years 2016–2019. The total number of papers included in the corpus is 158. The authors of these papers were all enrolled in master's or doctoral programs and had prior experience with academic paper writing. However, as their academic writing skills were in the early developmental stage, they can be considered novices in academic writing. Papers authored by individuals with experience in publishing papers in core journals were excluded through manual screening. Furthermore, papers that did not comply with the standards required for journal publication or could not be converted to the requisite format were excluded. The screening process yielded 135 papers that were retained for further analysis. To ensure comparability with the expert journal papers corpus, 95 papers were randomly selected from the 135 retained papers for analysis, amounting to a total of 336,000 words. The conference papers have been incorporated into the CNKI database. Although they have not yet been published in academic journals, the objective of these papers is consistent with that of journal papers, namely to facilitate academic discourse and exchange between peers. Moreover, the length of both conference papers and journal papers is comparable. Therefore, in addition to the discrepancy in paper quality, they are, for the most part, comparable. However, they differ significantly from dissertations in terms of both the purpose of the writing and the length of the texts. Accordingly, for the purpose of citation analysis of academic papers, we treat journal papers and conference papers as essentially equivalent and utilize the term "journal papers" to refer to both in our discourse.

3.2 Citation identification and coding

The identification of citation examples involves a systematic three-step process. The first step employs the HanLP toolkit, developed in Python, to perform Named Entity Recognition (NER) within the corpus, identifying entities such as names of individuals, places, and organizations. The second step consists of filtering out statements that reference the names of individuals. In the third step, statements are manually reviewed to exclude those that do not pertain to cited literature. Subsequently, the remaining statements are categorized according to the analytical framework outlined below.

In order to validate the generalisability of the findings on academic English citation practices among authors with different levels of expertise, this paper employs a citation example analysis framework inspired by Hyland (1999) research. The framework enables a comparison of the quantity of citations present in the papers of expert and novice authors. Furthermore, the study examines the citation practice of the two groups from three dimensions: embedding method, citation form, and the use of reporting markers.

Embedding method is classified into two categories: non-integral, as exemplified by (1), and integral. Integral can be divided into three categories depending on the syntactic position of the cited authors. The first category comprises instances where the cited author is the subject of the sentence, as illustrated in example (2). The second category encompasses cases where the cited author is not the subject but appears as an additional constituent, as exemplified in example (3). The third category includes instances where the cited author is in a noun phrase, as demonstrated in example (4).

- 1 The integration of Chinese culture teaching with Chinese language instruction has always been one of the important research topics in the field of Chinese language teaching (Lu and Ma, 2016).

中国文化教学与汉语语言教学的结合一直是汉语教学领域研究的重要课题之一(陆俭明, 马真2016).

- 2 Qiang (2010) distinguished between the topic marker “~嘛” and the modal particle “~嘛,” and described their process of grammaticalization.

强星娜(2010)则区分了话题标记的“~嘛”和语气词的“~嘛”, 并描写了它们语法化的过程。

- 3 As Mr. Lu pointed out, language teaching materials should not be confined to the systematic nature of linguistic knowledge when dealing with language materials...

正如鲁健骥先生所说, 语言教材在处理语言材料的时候, 不应拘泥于语言知识的系统性...

- 4 Among them, Sally's (2007) five insightful suggestions are as follows: first...

其中Sally (2007)提出的5条建议颇有见地...

Citation form is divided into four categories: block quote citation, which means that the quoted text is longer than or equal to 1 T unit, as illustrated in example (5); quote citation, which means that the original text is quoted as a word or phrase, as illustrated in example (6); summary citation, which means that the quoted text is a summary of one piece of literature, as illustrated in example (7); and generalization citation, which means that the quoted text is a summary of several pieces of literature, as illustrated in example (8). The above citation styles actually reflect the degree of integration of the original text by the author. Among them, block citation has the lowest level of integration of the original text, and the other three citation styles have increasing levels of integration, in that order.

- 5 The Ministry of Education's Department of Teacher Education... defines it as “the continuous development process of teachers as individual professionals, involving their continuous acquisition of new knowledge and enhancement of professional capabilities. To become a mature professional, teachers need to expand the depth of their profession and improve their professional level through continuous learning and exploration, thus achieving a state of professional maturity.” (Wang, 2015)

教育部师范教育司..., 将...界定为“教师个体专业不断发展的历程, 是教师不断接受新知识, 增长专业能力的过程。教师要成为一个成熟的专业人员, 需要通过不断的学习与探究历程来拓展其专业内涵, 提高专业水平, 从而达到专业成熟的境界”(王添淼 2015).

- 6 To eliminate the interference caused by relevant projects in language learning, George (1972) proposed an error prevention strategy called “orderliness of input,” suggesting that...

为了消除相关项目给语言学习带来的干扰, George (1972) 提出一种称作“有序输入”(orderliness of input) 的错误预防策略, 认为...

- 7 For example, “没”(méi) evolved from a verb to an adverb and gradually transformed into a subjectively diminishing marker due to the constraints of subjective expression (Zhang, 2006).

比如“没”从动词虚化为副词, 由于受到主观表达的制约逐渐转化成为主观减量标记(张谊生 2006).

- 8 The publication of the first set of Chinese textbooks for foreign language learners, “Chinese Textbooks,” in 1958 laid the foundation for... (Ke Bide, 1990; Li Quan and Jin Yunzhen, 2008, etc.)

1958年第一套对外汉语教材《汉语教科书》出版, 奠定了... (柯彼德 1990; 李泉, 金允贞 2008等) (徐晶凝 2016).

The selection of reporting markers reflects the rhetorical competence evident in academic writing. By selecting appropriate language forms and establishing intertextual relationships with external content, the writer is able to achieve the communicative purposes within the discourse.

The classification of reporting markers can be divided into three categories, depending on the criteria used for differentiation: research markers, cognitive markers, and discourse markers. Research markers are primarily associated with research acts and can be further classified into two subcategories: those pertaining to the research process and those pertaining to the results of the research. Those markers that refer to the research process, such as “examined” and “counted,” etc., and those that refer to the results of the research, such as “found” and “constructed,” etc. Cognitive markers, which mainly refer to cognitive processes, e.g., “concerned” and “speculated,” etc. Discourse markers, which refer mainly to speech acts such as “pointing out,” “elaborating,” etc.

According to Liu et al. (2021) and the corpus, the structural form of reporting markers in Chinese academic papers is very flexible, so we also examine the differences in the structural form of reporting markers between expert and novice papers. The structural form can be classified into four categories. The first category comprises independent verbs or independent verbs with a tense component, which are abbreviated as “v + le/quo.” Examples of this category include “propose,” “proposed.” The second category is a prepositional phrase, which is abbreviated as “pre + v.” An example of this category is “*dui...jinxing...yanjiu*.” The third category comprises reporting verbs situated in relational

TABLE 1 Citation practice analysis framework.

Embedding method	Non-integral, e.g., (Hyland, 1999)		
	Integral	Subject. As a subject, e.g., “Hyland (1999) argues that...”	
		Noun-phrase. The cited author is located in a noun phrase and the noun phrase serves as a necessary syntactic component, e.g., “Hyland (1999) study found that...”	
		Adjunct. The cited author is located in an adjunct phrase, e.g., “According to Hyland (1999)...”	
Citation form	Quote, citing only a word or phrase from the original text		
	Block quotes, e.g., Hyland (1999) suggests that “...”		
	Summary, e.g., Hyland (1999)		
	Generalization, e.g., “Hyland (1999) and Jiang, (2005)”		
Reporting markers	According to the concept of righteousness, divided into three categories	Research markers	Referring to the research process, e.g., “observe”
			Referring to the results of research, e.g., “find”
		Cognitive markers, e.g., “notice”	
		Discourse markers, e.g., “point out”	
	According to the structural form, divided into four categories	Reporting verbs alone or reporting verbs with an additional tense element, e.g., “v + <i>le/guo</i> ”	
		Prepositional phrase, e.g., “ <i>dui...jinxing le...v</i> ”	
		Reporting verbs are located in relational clauses, such as “...v <i>de n</i> ”	
		Reporting verbs are located in the central clause position of a modifier-head structure, e.g., “... <i>de v</i> ”	

clauses, which is abbreviated as “...v *de n*.” The fourth category encompasses the reporting verbs occupying the central clause position of a modifier-head structure, which is shortened to “...*de v*.”

In order to provide a more comprehensive overview of the analytical framework employed in this study, we have provided a summary of the aforementioned three categories in tabular form (see Table 1).

4 Results and discussion

4.1 Overall comparisons of citation practices across corpora

Table 2 presents the citation counts, average citations per paper, and relative citation rates for authors with varying degrees of expertise. The data indicate that expert authors demonstrate higher average citation counts and citations per thousand words compared to novice authors. However, the difference between the two groups is not statistically significant ($\chi^2 = 0.92, p > 0.05$). This finding aligns with previous research in academic English, which suggests that papers authored by individuals with higher levels of expertise tend to exhibit relatively higher citation rates (Mansourizadeh and Ahmad, 2011). Nevertheless, similar to the current study, the discrepancy between the two groups remains statistically insignificant (Li and Zhang, 2021).

The lower number of citations in novice authors’ papers may be attributed to a lesser degree of intertextuality awareness in this group. McCulloch (2013) conducted an exploratory analysis of the process undertaken by two master’s degree students from reading the material to writing a course paper, with a particular focus on the manner in which the authors utilized the source material to complete the paper. The study revealed that the level of intertextuality awareness

TABLE 2 Comparison of the number of citations by expert and novice authors.

Author Type	Number of citations	Average citations per paper	Relative citations (per 1,000 words)
Expert author	905	9.53	2.39
Novice author	765	8.05	2.28

exhibited by the authors varied considerably. Some of the authors demonstrated a higher degree of intertextuality awareness than others. This manifested in two ways. Initially, the authors demonstrated an active engagement with the source materials, extracting and adapting the information therein to express their own viewpoints. Secondly, they exhibited the ability to make connections between multiple source materials, extracting and utilizing the information after a critical comparison and reflection. In conclusion, authors with a high sense of intertextuality will consciously reshape information from source materials to apply it to their writing, and will actively expand and compare related materials for critical selection. Both of these behaviors can result in an increased number of discourse citations. It can therefore be surmised that the paucity of citations in the papers of novice authors is at least partly attributable to their limited awareness of intertextuality.

4.2 A comparison of expert and novice author citation practices

The number of citations in the papers of expert and novice authors is not significantly different. Nevertheless, this does not necessarily indicate that there are no discernible differences in the citation practices observed in Chinese academic papers between the two groups. The subsequent analysis will undertake a comprehensive

TABLE 3 Statistics on different types of author embedding methods.

Author Type	Non-Integral		Percentage	Integral		Percentage
	Raw numbers	Per 1,000 words		Raw numbers	Per 1,000 words	
Expert author	414	1.09	45.75%	491	1.3	54.25%
Novice author	268	0.8	35.03%	497	1.48	64.97%
Chi-square	$\chi^2 = 15.98, p < 0.05$			$\chi^2 = 4.19, p < 0.05$		

comparative examination based on the framework presented in Table 1.

4.2.1 Embedding method

A notable discrepancy was observed in the selection of the embedding method between the two author groups (see Table 3). Expert authors are more likely to utilize non-integral citations in comparison to their novice counterparts. As illustrated in Table 3, the proportion of non-integral citations among expert authors is 45.75%, whereas the corresponding figure for novice authors is 35.03%. This difference is statistically significant. This result is consistent with the findings of Mansourizadeh and Ahmad (2011), indicating that the observed effect of expertise on the choice of embedding style is generalizable across different linguistic contexts. Furthermore, academic Chinese exhibits distinctive characteristics with regard to the embedding method in comparison to English. In Chinese papers, there is a greater tendency toward integral than non-integral, whereas in English papers, the opposite is true.

There are three advantages to using non-integral citations as opposed to integrated ones. Primarily, situating the cited source outside of the sentence serves to accentuate the information contained within the citation, thereby facilitating a more objective presentation. Secondly, this approach enables authors to integrate the cited information seamlessly into their own viewpoint, thus making it an integral part of their argument. Thirdly, the use of non-integral citations ensures coherence within the discourse, preventing interruptions in the process of argumentation. These advantages of non-integral citations assist authors in developing their academic identities. In particular, the objective of introducing cited information is to construct the author's viewpoint, and non-integral citations are an effective means of achieving this goal. Authors construct their academic identities by forming their own perspectives based on the cited information and expressing them within the academic discourse community (Ma and Qin, 2015). The restricted deployment of non-integral citations by novice authors suggests a lack of awareness of the potential to actively shape their academic identities. Rather than critically reflecting on established perspectives to form their own unique viewpoints, their aim in incorporating cited information is often to seek the "correct answer" or to present existing viewpoints.

In accordance with the established analytical framework, there are three distinct syntactic positions for the cited authors in integrated citations. A comparison of the results reveals significant similarities in the syntactic positions of cited authors between the two types of authors (see Table 4). First, no notable discrepancy was identified between the two groups of authors in the syntactic positions occupied by the cited authors as subjects or within noun phrases. However, a notable discrepancy is evident when the cited authors are situated within an adjunct phrase. This result differs from the findings of

Mansourizadeh and Ahmad (2011), who observed a significant difference in the use of cited authors as subjects, with novice authors relying excessively on this structure (22.22%) compared to expert authors (6.56%). Second, both groups of authors demonstrate a markedly higher proclivity for utilizing cited authors as subjects, in comparison to the other two syntactic positions. The frequency of this pattern is markedly higher than the combined total of the other two patterns. This finding differs from that of Mansourizadeh and Ahmad (2011), in which the occurrence of cited authors as subjects was almost equal to the combined occurrence of the other two positions. Third, both groups of authors demonstrate a preference for utilizing cited authors in the following sequence: the preference for cited authors as subjects was observed to be the most frequent, followed by cited authors within an adjunct phrase and cited authors within a noun phrase. This observation is consistent with the conclusions of Arizavi and Choubsaz (2021), who conducted research on English-language academic journal papers and found that cited authors are most frequently placed as subjects, followed by prepositional phrases and noun phrases.

The preceding analysis indicates that the syntactic positions of cited authors in Chinese journal papers differ significantly from those in English papers. Nevertheless, the overall distribution pattern remains consistent with that observed in English papers. These discrepancies may be attributed to the distinctive characteristics of Chinese academic papers. In contrast to English papers, Chinese papers tend to place the cited authors in the subject position with greater frequency. This form is more accessible for novice authors, which may contribute to the absence of a significant difference between the two groups. Conversely, English papers frequently employ nominalized phrases, which may prove more challenging for those with limited writing experience and/or non-native proficiency. Some studies have demonstrated that non-native speakers utilize a reduced number of nominalizations in their written work in comparison to native speakers (Tambul ElMalik and Nesi, 2008). Consequently, novice authors frequently utilize citations with the cited authors in the subject position. With regard to the similarities, the disciplinary nature of linguistics may be the reason. Despite the differences between the two language corpora, they both belong to the same field of linguistics. The syntactic positioning of the cited authors may serve to illustrate the disparate value placed upon them by the authors in question. The differences in ontology, epistemology and methodology among disciplines result in varying emphases being placed on the source and the knowledge it represents. For example, applied linguistics tends to emphasize the authority of the source, whereas clinical psychology prioritizes the expertise of the knowledge acquisition process (Hu and Liu, 2020). The findings of this study indicate that both English and Chinese papers tend to cite authors in prominent subject positions, which can be attributed to the disciplinary nature of linguistics.

TABLE 4 Statistics on different types of author syntactic position.

Author type	Subject		Adjunct		Noun-phrase	
	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words
Expert author	364	0.96	91	0.24	36	0.09
Novice author	368	1.09	111	0.33	18	0.05
Chi-square	$\chi^2 = 3.01, p > 0.05$		$\chi^2 = 4.8, p < 0.05$		$\chi^2 = 3.52, p > 0.05$	

4.2.2 Citation form

A comparative analysis was conducted to examine the citation practices of expert and novice authors across four distinct categories. It was observed that, with the exception of quote, the two groups demonstrated notable discrepancies in their utilization of the remaining three citation forms, as illustrated in Table 5.

The discrepancies in the citation form between the two cohorts of authors can be encapsulated as follows: First, expert authors tend to employ a greater number of direct quotations in comparison to novice authors. This is demonstrated by the higher frequency of “block quote” and “quote” observed in the papers of expert authors in comparison to those of novice authors. This finding is consistent with the findings of Lombardi (2021), which also revealed an increase in the use of direct quotations with the writer’s level of expertise. Similarly, our study revealed that expert authors, who demonstrated greater proficiency, employed a greater number of direct citations than novice authors. The restricted deployment of direct quotations by novice authors indicates a diminished intertextual consciousness and affinity with source materials during the writing process. However, there is a discrepancy between the findings of our study and those of Lombardi (2021). While Lombardi (2021) observed a reduction in quotation length with increasing expertise levels, our study found that expert authors used longer “block quote” more frequently than novice authors did. We attribute this discrepancy to the differing nature of the corpora employed in each study. The papers of expert authors frequently comprise theoretical works that are heavily reliant on previous viewpoints. Consequently, it is imperative that they remain faithful to the original texts in order to guarantee the veracity of their arguments. Conversely, the papers of novice authors tend to comprise a greater proportion of content oriented toward application, which results in a lower incidence of opinion-based citations and a reduced necessity for extensive block quotations. Consequently, such citations are employed less frequently by novice authors.

A second distinction can be observed in the use of citations by expert and novice authors. Expert authors employ a greater number of generalization citations and a smaller number of summary citations compared to their novice counterparts. The utilization of generalization citations fulfils two distinct rhetorical functions. Primarily, it serves to enhance the credibility and authority of the content presented, thereby providing support for the author’s viewpoints or claims. This approach to citation enables authors to adapt the cited content in a flexible manner, thus enhancing the effectiveness of their argumentation and achieving the communicative goal of persuading readers (Hyland, 1999). Second, it establishes connections among numerous studies within the same field (Petrić, 2007), thereby demonstrating the author’s familiarity with the research domain and their ability to present themselves as an expert in academic writing. The restricted deployment of such citations by novice authors also suggests a deficiency in their intertextual

awareness with regard to existing research, as well as a lack of awareness of the selection of citation approaches that may be employed in order to construct an academic expert identity.

4.2.3 Reporting markers

A preliminary statistical analysis was conducted to examine the proportion of reporting markers utilized in citations and the frequency of high-frequency reporting verbs employed by the two groups of authors (see Table 6). It was observed that there were no significant differences in the frequency of reporting marker usage between the expert and novice authors.

The use of reporting verbs indicates that both groups of authors frequently utilize a similar set of high-frequency reporting verbs. However, expert authors demonstrate a higher level of diversity in their use of reporting verbs compared to novice authors, as evidenced by two key aspects. First, expert authors demonstrate a greater diversity of reporting verb types, resulting in a higher Type-Token Ratio (TTR) of reporting verbs in their corpus compared to novice authors. In particular, the TTR value for reporting verbs in the corpus of expert authors is 0.21, while in the corpus of novice authors, it is 0.2. Second, with regard to the coverage of high-frequency reporting verbs, the corpus of expert authors demonstrates that the top 10 high-frequency reporting verbs account for 53.11% of the total occurrences, whereas in the corpus of novice authors, the top 10 high-frequency reporting verbs cover 61.13% of the total occurrences. This suggests that novice authors tend to focus on utilizing the 10 most prevalent reporting verbs, exhibiting a lesser degree of complexity and adaptability in their paraphrase verb usage compared to expert authors. These findings are consistent with those of Lombardi (2021), which revealed that high-level authors exhibited a more diverse range of reporting verbs in their writing.

By analyzing the use of reporting markers with varying referential content, it is possible to ascertain the authors’ preferences with regard to the selection of original material. The comprehensive statistical findings are presented in Table 7. The distribution of the three types of markers is consistent in the corpora of both groups of authors. Discourse markers are the most frequently used, followed by research markers, and cognitive markers are the least used. However, there is a discernible discrepancy in the usage pattern between novice and expert authors. The data indicates that novice authors tend to utilize research process markers with greater frequency, while employing research result markers with lesser frequency, in comparison to expert authors. Lombardi (2021) observed that high-level authors tend to utilize reporting verbs that reflect their current discursive actions, such as “argue,” to express their evaluation of the cited content. In comparison to research process reporting markers, research result markers are more likely to convey evaluative information. To illustrate, the research result-oriented reporting marker “证实 (confirm)” indicates the author’s affirmative evaluation of the cited content. The

TABLE 5 Statistics on different types of author citation form.

Author type	Block quote		Quote		Summary		Generalization	
	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words
Expert author	178	0.47	104	0.27	476	1.26	147	0.39
Novice author	94	0.28	77	0.23	509	1.51	85	0.25
Chi-square	$\chi^2 = 16.43, p < 0.05$		$\chi^2 = 1.27, p > 0.05$		$\chi^2 = 8.45, p < 0.05$		$\chi^2 = 9.6, p < 0.05$	

TABLE 6 Percentage of reporting markers and high frequency reporting verbs.

Author type	Number of reporting markers		Percentage	High frequency reporting verbs (Frequency > 8)
	Raw numbers	Per 1,000 words		
Expert author	610	1.61	67.4%	point out (指出), think (认为), find (发现), say (说), propose (提出), study (研究), analyze (分析), explore (探讨), advocate (主张), examine (考察), classify (分为), 讨论 (discuss)
Novice author	566	1.68	73.99%	think (认为), propose (提出), point out (指出), analyze (分析), study (研究), examine (考察), mention (提到), explore (探讨), classify (分为), summarize (总结)
Chi-square	$\chi^2 = 0.55, p > 0.05$			

TABLE 7 Distribution of three types of reporting markers.

Author type	Research markers						Discourse markers		Cognitive markers	
	Research act		Research results		Total					
	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words
Expert author	171	0.45	59	0.16	230	0.61	365	0.96	15	0.04
Novice author	211	0.63	24	0.07	235	0.7	315	0.94	16	0.05
Chi-square	$\chi^2 = 10.07, p < 0.05$		$\chi^2 = 10.19, p < 0.05$		$\chi^2 = 2.19, p > 0.05$		$\chi^2 = 0.1, p > 0.05$		$\chi^2 = 0.11, p > 0.05$	

restricted deployment of research result reporting markers by novice authors indicates a deficiency in their ability to critically evaluate the cited information.

From a structural form perspective, there are significant differences in the use of the four types of structural forms between expert and novice authors (see Table 8). In particular, expert authors tend to favor the use of “v + *le/guo*” and “...*de* v,” while novice authors tend to use “pre + v.” The “...*de* v” structure serves two functions. First, it nominalizes the research process, making the expression more formal in writing style, as seen in example (9) with the word “调查” (investigation). Secondly, this structure provides syntactic positions for multiple paraphrased content. In example (9), it introduces the research object with the preposition “对” (regarding), and in example (10), it incorporates the manner information related to the reporting verb “倡导” (advocate) with the term “大力” (vigorously).

9 Tao Hongyin’s investigation of Chinatowns in the United States found that “compared with Mandarin in Hong Kong, Taiwan, and Southeast Asia, North American Chinese is more like a great fusion of Chinese varieties, as its users consist of immigrants from these diverse regions within the Chinese cultural sphere” (Li Yuming, 2017).

陶红印对美国唐人街的调查发现:“跟港台, 东南亚地区的华语相比, 北美汉语更像是一个汉语变体的大融合, 因为北美汉语使用者正是来自这些不同地区但同属中华文化圈的移民” (李宇明 2017).

10 Under the strong advocacy of Nattinger and De Carrico (1992), Lewis (1993, 1997), and others, the lexical approach, also known as “词汇法” in Chinese, has gradually become a influential teaching methodology.

在 Nattinger and De Carrico (1992), Lewis(1993, 1997) 等的大力倡导下, 语块教学法 (lexical approach,或译作“词汇法”) 逐渐成为一种较有影响的教学法流派.

The second function of this structure is to encapsulate the reporting information, allowing great flexibility in syntactic positioning and facilitating subsequent comments or evaluations. In example (11), the encapsulated information appears in the subject position and the author provides an evaluation of it. In example (12), the encapsulated information is placed in the object position, explaining the concept of “句本位” (sentence-based perspective). Similarly, in example (13), it also occupies the object position,

TABLE 8 Distribution of reporting markers' structural form.

Author type	v + le/guo		pre + v		...de v		...v de n	
	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words
Expert author	416	1.1	139	0.37	33	0.09	22	0.06
Novice author	302	0.9	244	0.73	5	0.01	15	0.04
Chi-square	$\chi^2 = 6.86, p < 0.05$		$\chi^2 = 42.25, p < 0.05$		$\chi^2 = 16.15, p < 0.05$		$\chi^2 = 0.39, p > 0.05$	

illustrating the basis of “教学呈现的先后顺序” (the order of instructional presentation), followed by further details of “主张” (proposition).

11 The analysis by Thomason is remarkably clear... (Zhang Bo, 2019)

Thomason的分析非常清楚,... (张博 2019).

12 Meanwhile, “句本位” is the proposition employed by Li Jinxi to elucidate the fundamental ideas of grammar (Zhao Jinming, 2017).

而“句本位”则是黎锦熙用以揭示语法基本思想的主张 (赵金铭 2017).

13 The sequence of presentation in teaching is based on Mr. Zhao Yuanren's proposition, primarily considering the frequency of phrase and structure usage (Zhao Jinming, 2018).

教学中呈现的先后顺序, 是依赵元任先生的主张, 主要考虑短语和结构的使用频率 (赵金铭 2018).

The “...de v” structure falls into the category of nominalization, which serves as a crucial “linguistic carrier” for conveying information in academic discourse (Gui, 2014, p. 51). The prevalence of this structure among expert authors indicates their ability to use language structures that are in line with academic discourse to reporting others' research and ultimately achieve their communicative goals.

There are differences in the temporal components attached to the reporting verbs used by the two groups of authors (see Table 9). Expert authors use “v + guo” more frequently and “v + le” less frequently compared to novice authors. Upon analyzing the corpus, we found that “进行了” (conducted) and “进行过” (have conducted) often alternate. To explore the differences in their usage, this study utilized Antconc 4.2.4 to examine high-strength collocates within the 8-word range to the right of both expressions. In the expert authors' corpus, the top 3 high-strength collocates for “进行了” are “统计” (count), “研究” (study), and “分析” (analyze), all of which belong to research reporting markers. On the other hand, the top 3 high-strength collocates for “进行过” are “论述” (discuss), “探讨” (explore), and “讨论” (discuss), which are all discourse reporting markers. The advantage of the “进行过 + discourse reporting marker” combination lies in its ability to provide an overall report of previous research, including but not

TABLE 9 Tense and aspect in reporting markers.

Author type	v + le		v + guo	
	Raw numbers	Per 1,000 words	Raw numbers	Per 1,000 words
Expert author	60	0.16	30	0.08
Novice author	139	0.41	8	0.02
Chi-square	$\chi^2 = 40.79, p < 0.05$		$\chi^2 = 9.26, p < 0.05$	

limited to the research process, with a stronger focus on the research results. This higher level of abstraction in the overall reporting allows the author to omit unnecessary reporting information, enabling them to emphasize their evaluation of previous research findings effectively.

5 Conclusion

Citations are an indispensable element of academic discourse, serving a pivotal function in the construction of knowledge, the interpretation of texts, and the dynamics of interpersonal communication. A substantial body of research on citations has been conducted, yielding a plethora of findings pertaining to various aspects of citations, including their forms, functions, and patterns across diverse contexts. Furthermore, differences in citation practices due to varying levels of expertise have been well-established in the field of academic English research. Nevertheless, it remains unclear whether the conclusions drawn from academic English corpora can be generalized to academic Chinese corpora. This study, which employed a self-constructed small-scale corpus, compared the similarities and differences in citation practices between expert and novice authors. The findings yielded three main conclusions:

First, it can be concluded that the findings derived from the analysis of academic English corpora can be largely extrapolated to academic Chinese corpora. This suggests that the impact of expertise on citation practices in academic journal papers is a cross-linguistic phenomenon. This study demonstrates that in academic Chinese writing, expert authors and novice authors exhibit comparable differences in citation density, embedding methods, citation forms, and reporting markers, as observed in academic English. For instance, expert authors are more likely to utilize non-integral citations and direct quotations, employ a diverse array of reporting markers and exhibit discernible proclivities in the utilization of evaluative reporting markers.

Moreover, academic Chinese exhibits distinctive features that set it apart from academic English. In academic Chinese papers, authors demonstrate a greater proclivity for employing the form of situating citees in subject position in comparison to their counterparts in the field of academic English. This form is readily comprehensible, resulting in no notable discrepancies in its utilization between expert and novice authors. Conversely, academic English writing tends to favor nominalized phrases, with the cited author situated within a noun phrase. This structure can prove challenging for novice authors to master, leading to an overreliance on the citees as subject form and resulting in significant usage differences between the two groups in this regard.

Ultimately, the discrepancies in citation practices between the two cohorts of authors can be attributed to their comparatively weaker intertextual awareness and less pronounced sense of developing an academic writing expert identity. In particular, novice authors tend to introduce cited information with the objective of identifying the “correct answer,” rather than engaging in a critical integration of disparate pieces of information and establishing intertextual relationships between the current discourse and multiple source materials. This approach fails to demonstrate their expertise or construct an expert identity in academic writing. The latter is achieved through synthesizing various sources, forming their own academic perspectives and highlighting their professional knowledge.

Two limitations remain in this study. First, the analytical framework addresses the form of citation, but not the function of citation. A combination of formal and functional analyses would have enabled the formulation of more operational pedagogical suggestions and provided novice authors with a clearer understanding of the appropriate citation forms for fulfilling communicative purposes. Second, the analysis is confined to the textual corpus; however, if interviews with novice and expert authors were to be incorporated, the motivations behind the observed differences in citation use between the two groups could be subjected to more rigorous analysis, thereby enhancing the reliability of the conclusions drawn.

References

- Ahn, C.-Y., and Oh, S.-Y. (2024). Citation practices in applied linguistics: a comparative study of Korean master's theses and research articles. *J. Engl. Acad. Purp.* 69:101369. doi: 10.1016/j.jeap.2024.101369
- Arizavi, S., and Choubasaz, Y. (2021). Citation practices in research article introductions: the interplay between disciplines and research methodologies. *Int. J. Appl. Linguist.* 31, 383–405. doi: 10.1111/ijal.12337
- Bazerman, C. (1988). *Shaping written knowledge*. Madison, WI: University of Wisconsin Press.
- Borg, E. (2000). “Citation practices in academic writing” in *Pattern and perspectives: Insights into EAP writing practice*. ed. P. Thompson (Reading: Centre for Applied Language Studies, University of Reading), 26–44.
- Charles, M. (2006). Phraseological patterns in reporting clauses used in citation: a corpus-based study of theses in two disciplines. *Engl. Specif. Purp.* 25, 310–331. doi: 10.1016/j.esp.2005.05.003
- Chen, C., and Zhang, L. J. (2017). An intercultural analysis of the use of hedging by Chinese and Anglophone academic English writers. *Appl. Ling. Rev.* 8, 1–34. doi: 10.1515/applrev-2016-2009
- Cui, L., and Cheng, X. (2014). Contrastive study of the applications of verbal evidentials in Chinese and English academic papers. *J. Dalian Univ. Technol.* 35, 115–119. doi: 10.19525/j.issn1008-407x.2014.02.021
- Fazel, I., and Shi, L. (2015). Citation behaviors of graduate students in grant proposal writing. *J. Engl. Acad. Purp.* 20, 203–214. doi: 10.1016/j.jeap.2015.10.002
- Gui, S. (2014). *Corpus-based analysis of the linguistic genre of English for linguistics*. Beijing: Foreign Language Teaching and Research Press.
- Harwood, N. (2009). An interview-based study of the functions of citations in academic writing across two disciplines. *J. Pragmat.* 41, 497–518. doi: 10.1016/j.pragma.2008.06.001
- Hu, G., and Liu, Y. (2020). Disciplinarity and academic discourse. *Foreign Lang. Educ.* 41, 29–33. doi: 10.16362/j.cnki.cn61-1023/h.2020.02.006
- Hu, G., and Wang, G. (2014). Disciplinary and ethnolinguistic influences on citation in research articles. *J. Engl. Acad. Purp.* 14, 14–28. doi: 10.1016/j.jeap.2013.11.001
- Hyland, K. (1999). Academic attribution: citation and the construction of disciplinary knowledge. *Appl. Linguis.* 20, 341–367. doi: 10.1093/applin/20.3.341
- Hyland, K. (2010). Metadiscourse: mapping interactions in academic writing. *Nordic J. Eng. Stu.* 9, 125–143. doi: 10.35360/njes.220
- Hyland, K., and Jiang, F. (2019). Points of reference: changing patterns of academic citation. *Appl. Linguis.* 40, 64–85. doi: 10.1093/applin/amx012
- Kwan, B. S. C., and Chan, H. (2014). An investigation of source use in the results and the closing sections of empirical articles in information systems: in search of a functional-semantic citation typology for pedagogical purposes. *J. Engl. Acad. Purp.* 14, 29–47. doi: 10.1016/j.jeap.2013.11.004
- Lee, J. J., Hitchcock, C., and Elliott Casal, J. (2018). Citation practices of L2 university students in first-year writing: form, function, and stance. *J. Engl. Acad. Purp.* 33, 1–11. doi: 10.1016/j.jeap.2018.01.001
- Li, J. (2011). *Contrastive genre analysis of the Chinese and English introduction sections of research articles in applied linguistics*. Shanghai: Shanghai International Studies University.
- Li, X. (2012). *Exploring the generic features of MA theses in linguistics*. Zhejiang: Zhejiang University.

Author contributions

XG: Writing – original draft, Writing – review & editing. RL: Writing – original draft, Writing – review & editing. CJ: Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was financially supported by the National Social Science Fund of China under the project ‘Corpus-based research on linguistic features of academic Chinese’ (NSSF Grant No. 20BYJ120).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Li, Q., and Zhang, X. (2021). An analysis of citations in Chinese English-major Master's theses and doctoral dissertations. *J. Engl. Acad. Purp.* 51:100982. doi: 10.1016/j.jeap.2021.100982
- Liu, R., Huang, Q., and Wang, S. (2021). The form, function and distribution of reporting markers in Chinese academic papers. *Contempor. Rhetoric* 6, 60–72. doi: 10.16027/j.cnki.cn31-2043/h.2021.06.007
- Lombardi, A. (2021). More is more: explicit intertextuality in university writing placement exam essays. *J. Engl. Acad. Purp.* 50:100955. doi: 10.1016/j.jeap.2020.100955
- Lou, B. (2011). Reporting verbs in master's theses of English learners. *Foreign Lang. Bimonthly* 34, 64–68.
- Ma, R., and Qin, X. (2015). A review of research on the use of literary resources for second language academic writing abroad. *Foreign Lang. Bimonthly* 38, 85–92.
- Mansourizadeh, K., and Ahmad, U. K. (2011). Citation practices among non-native expert and novice scientific authors. *J. Engl. Acad. Purp.* 10, 152–161. doi: 10.1016/j.jeap.2011.03.004
- Marti, L., Yilmaz, S., and Bayyurt, Y. (2019). Reporting research in applied linguistics: the role of nativeness and expertise. *J. Engl. Acad. Purp.* 40, 98–114. doi: 10.1016/j.jeap.2019.05.005
- Martínez, I. A. (2008). Building consensus in science: resources for intertextual dialog in biology research articles. *J. Engl. Acad. Purp.* 7, 268–276. doi: 10.1016/j.jeap.2008.10.011
- McCulloch, S. (2013). Investigating the reading-to-write processes and source use of L2 postgraduate students in real-life academic tasks: an exploratory study. *J. Engl. Acad. Purp.* 12, 136–147. doi: 10.1016/j.jeap.2012.11.009
- Mu, C. (2024). Citation choices in L2 novices' and experts' literature review sections: a functional discourse analysis. *J. Engl. Acad. Purp.* 68:101361. doi: 10.1016/j.jeap.2024.101361
- Peng, J.-E. (2019). Authorial voice constructed in citation in literature reviews of doctoral theses: variations across training contexts. *J. Engl. Acad. Purp.* 37, 11–21. doi: 10.1016/j.jeap.2018.11.001
- Petrić, B. (2007). Rhetorical functions of citations in high- and low-rated master's theses. *J. Engl. Acad. Purp.* 6, 238–253. doi: 10.1016/j.jeap.2007.09.002
- Petrić, B. (2012). Legitimate textual borrowing: direct quotation in L2 student writing. *J. Second. Lang. Writ.* 21, 102–117. doi: 10.1016/j.jslw.2012.03.005
- Samraj, B. (2013). Form and function of citations in discussion sections of master's theses and research articles. *J. Engl. Acad. Purp.* 12, 299–310. doi: 10.1016/j.jeap.2013.09.001
- Shi, W. (2013). Intentional patterns of reporting clauses in research articles by Chinese authors: a Corpus-based study. *Foreign Lang. Educ.* 34, 27–31. doi: 10.16362/j.cnki.cn61-1023/h.2013.03.012
- Sun, Y. (2009). An analysis of citations used in introductions of Chinese MA thesis. *Foreign Lang. Educ.* 30, 53–69. doi: 10.16362/j.cnki.cn61-1023/h.2009.01.018
- Swales, J. M. (1986). Citation analysis and discourse analysis. *Appl. Linguis.* 7, 39–56. doi: 10.1093/applin/7.1.39
- Swales, J. M. (2014). Variation in citational practice in a Corpus of student biology papers: from parenthetical plonking to intertextual storytelling. *Writ. Commun.* 31, 118–141. doi: 10.1177/0741088313515166
- Tambul ElMalik, A., and Nesi, H. (2008). Publishing research in a second language: the case of Sudanese contributors to international medical journals. *J. Engl. Acad. Purp.* 7, 87–96. doi: 10.1016/j.jeap.2008.02.007
- Thompson, P., and Tribble, C. (2001). Looking at citations: using corpora in English for academic purposes. *Lang. Learn. Technol.* 5:91.
- Thompson, G., and Ye, Y. (1991). Evaluation in the reporting verbs used in academic papers. *Appl. Linguis.* 12, 365–382. doi: 10.1093/applin/12.4.365
- Wang, G., and Hu, G. (2022). Citations and the nature of cited sources: a cross-disciplinary and cross-linguistic study. *SAGE Open* 12:21582440221093350. doi: 10.1177/21582440221093350
- Wette, R. (2018). Source-based writing in a health sciences essay: year 1 students' perceptions, abilities and strategies. *J. Engl. Acad. Purp.* 36, 61–75. doi: 10.1016/j.jeap.2018.09.006
- White, H. D. (2004). Citation analysis and discourse analysis revisited. *Appl. Linguis.* 25, 89–116. doi: 10.1093/applin/25.1.89
- Zhang, G. (2022). The citational practice of social science research articles: an analysis by part-genres. *J. Engl. Acad. Purp.* 55:101076. doi: 10.1016/j.jeap.2021.101076
- Zhang, G. (2023). Authorial stance in citations: variation by writer expertise and research article part-genres. *Engl. Specif. Purp.* 70, 131–147. doi: 10.1016/j.esp.2022.12.002
- Zhao, H., and Zhan, H. (2020). The use of genitive citations in academic writing. *Lingua* 234:102769. doi: 10.1016/j.lingua.2019.102769



OPEN ACCESS

EDITED BY

Xinghua Liu,
Shanghai Jiao Tong University, China

REVIEWED BY

Carol Nash,
University of Toronto, Canada
Hariharan N. Krishnasamy,
Universiti Utara Malaysia, Malaysia
Barry Lee Reynolds,
University of Macau, China
Alfonso Garcia De La Vega,
Autonomous University of Madrid, Spain

*CORRESPONDENCE

Sandra Becker
✉ sandra.becker@ucalgary.ca

RECEIVED 01 November 2024

ACCEPTED 07 April 2025

PUBLISHED 25 April 2025

CITATION

Becker S, Jacobsen M and Friesen S (2025)
Four supervisory mentoring practices that
support online doctoral students' academic
writing.
Front. Educ. 10:1521452.
doi: 10.3389/feduc.2025.1521452

COPYRIGHT

© 2025 Becker, Jacobsen and Friesen. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Four supervisory mentoring practices that support online doctoral students' academic writing

Sandra Becker*, Michele Jacobsen and Sharon Friesen

Werklund School of Education, University of Calgary, Calgary, AB, Canada

Academic writing in both face-to-face and online environments is often fraught with tension, emotion, and challenge. The quality of doctoral students' online academic writing experiences can be a difference maker in the successful completion of programs. This study examines how mentoring practices support online doctoral students' academic writing, building on prior research that identified five enabling factors of effective online doctoral supervision, with a focus on cultivating a collaborative online community of support for academic writing. Using a comparative case study approach, interviews with five recently completed faculty of education doctoral graduates at a large university in western Canada were analyzed to identify four mentoring supervisory practices that support online doctoral students' academic writing: (a) fostering a trusting, supportive community of practice; (b) engaging in regular synchronous meetings combined with iterative cycles of mentoring and scaffolding; (c) using coursework and program structures as a springboard for writing; and (d) providing diverse models of academic writing. Central to the effectiveness of the four online supervisory mentoring practices was the notion of trust which enabled students to develop their academic writing skills, scholarly identities, and successfully complete their doctoral degrees. This study is significant for identifying supervisory mentoring practices that led to students' sense of gratitude and flourishing, further highlighting how crucial relational trust is for online doctoral students' academic writing.

KEYWORDS

Online doctoral supervision, academic writing, supervisory practices, graduate mentorship, graduate student writing

Introduction

Doctoral work is often inspired by an innate sense of curiosity, an array of life experiences and diverse personal passions, and the joy found in learning new things. A substantial component of doctoral work involves communication, and particularly the complex and challenging process of scholarly writing (Calle-Arango and Ávila Reyes, 2023; Naidoo et al., 2023; Ondrusek, 2012). Seasoned scholars often repress or downplay the emotional demands and vulnerabilities associated with academic writing (Belcher, 2019; Goodson, 2023), yet these challenges remain significant barriers, particularly for doctoral students navigating the transition to academic writing. Challenges with academic writing have implications for doctoral student-supervisor relationships, such as communication, feedback, and student confidence. Exploring first time writing and feedback experiences in a doctoral program, Wei et al. (2019) identified the shift from the undergraduate and master's level—where students encountered a high degree of success and limited critical commentary—to doctoral programs which included highly evaluative peer review and critique, and expectations for not only the

sheer quantity of writing, but also multiple rounds of revision required for each work. Further, although there are beliefs within faculty that students enter the doctoral program able to write, many students express a lack of preparedness (Calle-Arango and Ávila Reyes, 2023). Indeed, transitioning from “coursework-based study to a doctoral degree involves a significant move to unstructured study and the requirement to produce a high-quality academic product contributing new knowledge to a field of study” (Bastalich and McCulloch, 2022, p. 1).

Mentoring practices that support doctoral students in making transitions to advanced academic writing are key to effective student-supervisor relationships, whether these interactions are mediated face-to-face or online. While there is a growing body of research focusing on specific interventions and strategies such as writing conferences (Consalvo and Rueter, 2024), group activities such as workshops, snack writing, and writing commons (Eaton and Dombroski, 2022; Maldonado et al., 2021; Sarnecka et al., 2022; Winberg et al., 2023; Zimmer et al., 2022), these studies primarily focus on short-term interventions rather than the overarching supervisory mentoring practices that sustain long-term writing development. This study addresses these gaps by exploring two key research questions: (1) What supervisory mentoring practices are identified as cultivating effective online doctoral student academic writing relationships? and (2) What emotions are associated with effective online doctoral student academic writing relationships?

In online environments, academic writing introduces unique challenges, such as the need for structured and timely communication with supervisors, as well as prompt and constructive feedback (Jacobsen et al., 2021; Naidoo et al., 2023). Research has shown that fostering a sense of belonging and writing self-efficacy plays a critical role in doctoral students overcoming challenges, emphasizing the importance of trust and community within virtual doctoral programs (Burkholder and Bidjerano, 2023; Miller et al., 2023). While existing research highlights various interventions to support academic writing (Maldonado et al., 2021; Sarnecka et al., 2022), there remains a need to explore the practices that integrate these strategies into cohesive and effective supervisory frameworks.

This study builds on earlier research identifying five enabling factors of effective online doctoral supervision, with a specific focus on factor five: cultivating a collaborative online community of support for academic writing (Jacobsen et al., 2021). Using a comparative case study approach (Merriam and Tisdell, 2016; Cleland et al., 2021), we analyzed data from interviews with five recently completed doctoral graduates at a large research-intensive Canadian university to determine the mentoring practices associated with effective online doctoral student academic writing skills. Findings include four supervisory mentoring practices: (a) engaging students in a trusting, supportive community of practice; (b) engaging in regular synchronous meetings combined with iterative cycles of mentoring and scaffolding; (c) using coursework and program structures as a springboard for writing; and (d) providing diverse models of academic writing. Central to the effectiveness of these online mentoring practices was the notion of trust.

A qualitative case study research design (Merriam and Tisdell, 2016) allowed for an in-depth exploration of the supervisory mentoring practices that supported online doctoral students' academic writing. While much of the existing research on doctoral writing focuses on short-term interventions (e.g., workshops, writing groups),

this study examined sustained supervisory mentoring as a relational and developmental practice that extends across a doctoral student's program. The doctoral graduates indicated that these effective supervisory mentoring practices not only assisted them in developing their academic writing practice as a tool for communication, thinking, creating new knowledge, and forming their academic identities, it ultimately lead to a deep sense of gratitude upon the completion of their doctoral degree. By explicitly addressing the identified research gap with a focus on the mentoring practices that supported student success, this case study advances the understanding of how relational trust operates in online supervision, shaping not only academic writing development but also students' confidence and scholarly identity. This study also highlights the emotional dimensions of online doctoral writing, demonstrating how supervisory mentorship contributes to student flourishing. The significance of this study lies in its contribution to improving supervisory mentorship in online doctoral education by underscoring the importance of relational trust in intentional supervisory mentorship in online doctoral programs to enhance student confidence and flourishing in academic writing.

Doctoral student writing experiences

Doctoral writing experiences in general are often emotionally fraught (Adamek, 2015; Colombo, 2018; Everitt, 2022; Huerta et al., 2017; Lavelle and Bushrow, 2007; Kirkpatrick, 2019) and include emotional responses such as fear, trepidation, stress, anxiety, feelings of low self-confidence and being overwhelmed. We argue that supervisory attention to writing as a key aspect of the online doctoral experience ensures that students develop a growing sense of belonging and self-efficacy in the academy. Supervisors, however, are not only confronted with *student* feelings in relation to the writing process, professors are also dealing with their own emotions affiliated with positive and negative writing experiences as they mentor doctoral student's academic progress (Belcher, 2019; Goodson, 2023). The rational-emotional combination presents a double helix predicament when considering supervisory practices and support for doctoral writing. In addition, supervisors need to be aware of different approaches that students may bring to the writing process (Berdanier, 2021; Lavelle and Bushrow, 2007). Lavelle and Bushrow for example, describe different doctoral student writing styles such as task-oriented (get it over with), systems-oriented (big picture), structure-oriented (organized), and intuitive (deep).

Academic writing is considered by some to be the most challenging aspect of graduate work; thus, supportive mentoring practices are needed to address the emotional, cognitive, conceptual, and methodological processes students undergo (Calle-Arango and Ávila Reyes, 2023; Colombo, 2018; Stevens and Caskey, 2023). It is not just about the writing per se - the affective, theoretical, and methodological aspects of academic composition can institute roadblocks to the construction of ideas. The notion of roadblocks aligns well with Belcher's (2019) suggestion to focus on, when writing, the large-scale aspects of the article - “its argument, evidence, structure, findings or methods” (p. 204). Furthermore, student scholars identify the need for clear expectations (Wei et al., 2019; Stevens and Caskey, 2023) as well as the provision of experiences and support in the giving and receiving of feedback (Carter et al., 2020; Catterall et al., 2011; Calle-Arango and Ávila Reyes, 2023; Chakraborty

et al., 2021; Kirkpatrick, 2019; Kumar and Coe, 2017; Larcombe et al., 2007; Wei et al., 2019). In addition, program structures must consider the inclusion of structured writing practice as part of course work (Stevens and Caskey, 2023; Larcombe et al., 2007) whereby assignments are seen as “springboard[s]” (Simpson, 2012, p. 106) to less structured, candidate-led academic writing for the dissertation (Hutton et al., 2024).

Doctoral writing in online environments introduces unique challenges, such as the need for structured and timely communication with supervisors along with prompt constructive feedback (Jacobsen et al., 2021; Can and Walker, 2014). Though research on the complexities of online doctoral writing is “sparse but growing” (Kirkpatrick, 2019, p. 22; Lee et al., 2024), studies that explore students’ experiences with online academic writing are essential to deepen understanding of effective supervision and mentoring practices that support doctoral students. In our previous research, both supervisors and doctoral graduates surfaced specific strategies, such as ensuring regularly scheduled productive meetings, and the use of a variety of online communication tools and spaces (Google docs, Zoom, text messages, emails, telephone calls) for staying connected to ensure the required responsiveness (Jacobsen et al., 2021).

In summary, the complex nature of doctoral student writing, both emotive and intellectual, whether in person or online, necessitates the need for ongoing support to ensure the successful completion of the doctoral degree. This support should be multifaceted, including regular, structured feedback, explicit guidance on academic writing expectations, and sustained mentoring practices that recognize the affective dimensions of writing. Establishing relational trust between students and supervisors is particularly critical, as this personal regard for one another fosters an environment where students feel supported in their academic identity formation and writing development.

Exploring notions of trust in doctoral student-supervisory relationships

A recent study (Jacobsen et al., 2021) identified five enabling factors key to fostering strong online student supervisor relationships with factor five being, “Cultivating a collaborative community of support for academic writing” (p. 3). Each of the five enabling factors were present in successful student supervisor relationships, however, the meta-factor across all was relational trust.

Etymologically, the *Oxford English Dictionary* (2024) indicates the origins of the word trust are Germanic, meaning “the state or condition of having something committed to one’s care or safekeeping; or of having confidence or faith placed in one; guardianship,” with “loyalty, reliability, trustworthiness” also referenced. At the heart of the word trust is a solid commitment to be bound with one another in honesty, reliability, and support.

Research on supervisory relationships demonstrates that relational trust is key to ensure that online supervisor-student relationships thrive (Jacobsen et al., 2021; Friesen et al., 2022). Relational trust evolves with both student and supervisor investing time in effective communication, and perspective taking, that starts early in the relationship and can lead to confidence in the judgements of each other. Honest communication, which starts with the supervisor, is necessary for both participants to flourish during the inevitable ups and downs of the doctoral journey, especially given the unequal power

dynamic. The supervisor is obligated to establish a trustworthy environment where the student feels safe in expressing themselves, even when this action requires great vulnerability (Makhamreh and Kutsyuruba, 2021). Trust, developed reciprocally over time, can easily be forfeited unless students witness consistency of action in their supervisor, signalling to them that they have their “best interests at heart” (Makhamreh and Kutsyuruba, 2021, p. 129). Consistency of action is often shown through constructive feedback, informed guidance, and open dialogue (Jacobsen et al., 2021; Friesen et al., 2022), which can sometimes be more challenging to maintain in online settings.

In summary, relational trust is fundamental to successful online doctoral supervision, serving as the foundation for strong student-supervisor relationships. Relational trust fosters reciprocity, consistency, and open communication, allowing students to feel safe in expressing academic vulnerabilities and developing scholarly identities. Given the inherent power imbalance in supervision, trust must be intentionally cultivated and sustained through honest dialogue, constructive feedback, and consistent mentorship.

Method

Research design

This research is derived from qualitative (Merriam and Tisdell, 2016) and explanatory case study research (Cleland et al., 2021; Yin, 2018) that identified five key enabling factors in effective online doctoral supervision (Jacobsen et al., 2021). The bounded system includes five recently completed doctoral graduates and five supervisors, from a large, research intensive university in western Canada who were purposely selected, invited, and agreed to be interviewed regarding their online supervisory relationships and experiences based on the recent, successful completion of their doctoral degree. All doctoral graduate participants were experienced educators (K-16) who had completed their doctoral studies while working in professional situations as educators and/or administrators. Of the five participants, four were female and all were first language English speakers. The study was approved by the Conjoint Faculties Ethics Review Board (CFREB). All participants provided informed consent prior to enrolment in the study.

Data collection

In the present analysis, we focus specifically on the semi-structured one-hour interviews with five doctoral graduates to examine in-depth the mentoring practices they associated with online academic writing experiences in graduate school. We identified ‘cultivating a community of support for online academic writing’ as a key enabling factor in the broader study. Building on this one factor, our secondary analysis focused on understanding the supervisory mentoring practices that doctoral graduates recognized as instrumental to developing their writing abilities.

Interview data was anonymized prior to analysis; participants were only identifiable as doctoral graduates who successfully defended their dissertation. Names used in reporting are pseudonyms.

Data analysis

A first cycle of deductive coding was conducted by all three authors who read each of the five transcripts to identify statements related to academic writing, then coding these chunks of text using the five enabling factors (Jacobsen et al., 2021) derived from our prior research. The data was organized in a spreadsheet, and we did a frequency count of codes to identify and rank the most prevalent to the least prevalent factor.

A second round of emotive coding provided insights into participant perspectives and experiences with an explicit focus on the emotions associated with academic writing (Saldaña, 2021). Inductive or open coding of emotions expressed in reference to academic writing was derived from the language used by participants.

Next, we reviewed the data within each enabling factor to identify patterns that could be collapsed into categories or themes related to graduates' descriptions of supervision and mentoring practices associated with academic writing. The resulting themes were then compared with themes emerging from other factors and the extant literature. We added direct quotes from the data into the coding framework to assist with the categorization. Interrater reliability was maintained through consensus building, using a process of reflexive thematic analysis (Braun and Clarke, 2019, 2020; Braun et al., 2022).

Findings

The findings are presented in two sections, with each section related to one of the research questions.

Research Question One: What supervisory mentoring practices are identified as cultivating effective online doctoral student academic writing relationships?

We pinpointed four mentoring practices identified by doctoral students, that working in tandem, served as instrumental to their success in online academic writing.

Practice 1: engaging students in a trusting and supportive collaborative community of practice

An important practice participants indicated as key to their progress, and linked to our previous research on relational trust, was the trusting and collaborative online community environment, one that was created based on the values of reciprocity, genuine caring, integrity, and respect (Bryk and Schneider, 2002; Friesen et al., 2022). As Alex attempted to articulate, "You cannot, you do not, it's not stand alone." The supervisor was often at the forefront in facilitating the creation and sustainment of this trusting online community which included the supervisory committee, course instructors, and student peers in the program.

According to some participants, the supervisory committee played a significant role and was important within the collaborative community of practice, with all members working as a team to support the doctoral student's writing. This team approach to mentoring academic writing yielded multiple benefits, such as

providing ready access to diverse faculty members who possessed a broad background and experience where their "strengths were different from my advisor's strengths" (Alex). The committee provided students with multiple experts they could trust and turn to regarding disciplinary knowledge and methodological approaches to research, data analysis, and the drafting of manuscripts. The supervisor, in helping to select committee members, and in leading mentoring conversations with the committee during online meetings, played an important leadership role in establishing this collaborative community of support. The committee also filled in when the supervisor was unable. For example, Shawn stated, "[Supervisor] got really busy with work. However, [Supervisor] had scaffolded it so other people were there to support me."

In addition to regular online connections with their supervisor and committee, graduates emphasized the value of their online cohort as a community of practice and source of peer mentorship. Peers served as critical friends who offered authentic feedback and diverse perspectives, while also questioning and challenging each other. This online community of peers was important because together the doctoral students were learning how to *be* academics by engaging in thoughtful critiques of not only their own ideas, but of the ideas expressed by peers. Peer support meant the online cohort engaged in a shared experience where all ideas were valued, and where the doctoral students motivated each other to keep going and care for each other. In describing their experience, Morgan stated

Our cohort had developed some pretty tight relationships. And so the feedback was more authentic, I would say, more true critical friend, where, I care about you, you have great research ideas, and I'm going to give you feedback that's going to nudge you, not just help you feel good about yourself.

Morgan highlighted the role of peer support in helping doctoral students not only improve their work, but also build a sense of belonging and motivation within the academic community.

Alex described how when supervisory scaffolding was lacking, they sought support from their cohort and supervisory committee, whom they described as their "life preserver." Alex compared the depth of feedback received from their supervisor to the committee members:

Whereas they [the Committee] were like, what are you actually saying in this? My advisor was more around the wordsmithing of stuff and you know, cropping things, so I was just stating as opposed to telling a story, they [the committee] were kind of like what is the information you're putting out there and what are you trying to get people to receive from this and how should you structure it so that it really has an impact?

The combination of supervisory, peer, and committee feedback created a comprehensive support system that fostered the academic growth of these doctoral students, providing them with multiple avenues to refine their ideas and writing.

Alex's statement speaks to the challenge for doctoral students where writing is used as a mechanism for constructing knowledge. Scholars acknowledge the complex process of writing to create new knowledge that moves beyond the telling of ideas to transformation, which requires support (Ondrusek, 2012; Jackman et al., 2024). In

Alex's case, this support came not from the supervisor, but from the supervisory committee and their peer cohort.

I don't want it to sound like I didn't have support because I did. Every time I emailed my advisor they got back to me, but I got a lot of my support in terms of writing my dissertation from two people that were in my cohort. I got most of my support from writing the actual, the content and how it should flow, from them.

This reliance on peers and committee members highlights the importance of multiple layers of support for doctoral students, ensuring they have access to guidance that meets their academic and emotional needs throughout the dissertation writing process.

Alex's metaphor of "life preserver," as committee support in a very real sense, spoke to their feelings of abandonment by the supervisor and rescue by committee members and peers. Although their supervisor's email replies signalled an acknowledged form of "support," what Alex seemed to be longing for as indicated by her comments about the supervisory committee, was a trusting and relational dialogue around the expression of ideas they were exploring.

Other participants recounted that the most important relationship in their online doctoral program was their supervisor. As Jesse explained, "I had a really good supervisor who was willing to work with me in a way that made sense for us. And that made the program better for me." Although Jesse indicated they dialogued with members of their cohort about general education topics, they did not explore their writing and writing process with peers in depth. "I mean, other than talking about education, just generally what people were thinking about, it really felt like those were two separate processes." In essence, for Jesse, the writing support and feedback was provided solely by their supervisor.

When it came to writing the dissertation, Morgan expressed a longing for associations beyond their supervisor and committee. "With my program, I did not have enough of a connection with other people outside of my supervisor and my supervisory committee at the back end," suggesting the benefit of creating a trusting *community* of support that exists beyond the supervisor and committee that students can draw upon as needed. The importance of a community of support for doctoral writing was identified in our larger study (Jacobsen et al., 2021) and is echoed in other research (Kar, 2024; Kirkpatrick, 2019; Wikeley and Muschamp, 2004).

The supervisor, in possessing more intimate knowledge of the doctoral student, however, can make the difference in a student thriving instead of merely surviving the academic writing journey. The strong relationships that four participants, Morgan, Shawn, Jesse, and Leslie described with their supervisors linked to a balanced connection to a broader community of practice in which they thrived. For Alex, a weak relationship with their supervisor meant they experienced abandonment and isolation and had to draw heavily on the extended community of support to survive. Alex described a lack of confidence while completing their dissertation and for undertaking further research, while highlighting the need for a competent and responsive supervisor.

The combination of supervisory, peer, and committee engagement and feedback created a comprehensive support system that fostered the academic growth of these doctoral students, providing them with multiple avenues to refine their ideas and writing. However, addressing the challenges and barriers to academic writing, including emotional

investment and vulnerability, cannot fall entirely on the supervisor. Establishing a collaborative community of practice that actively engages in supportive measures and ongoing feedback is an important consideration in program and institutional structures (Adamek, 2015; Aitchison and Paré, 2012; Belcher, 2019; Catterall et al., 2011; Chakraborty et al., 2021; Stevens and Caskey, 2023), while "resituat[ing] the supervisor, not as the sole provider of support but as an important anchor within a network of wider contacts and supports" (Bastalich and McCulloch, 2022, p. 9).

Practice 2: engaging in regular synchronous meetings with iterative cycles of mentoring and scaffolding

The most prevalent supervisor practice participants described for cultivating productive online writing relationships was regular and iterative cycles of mentoring and scaffolding. This practice occurred during various phases of written doctoral work: (i) during coursework and dissertation background reading and pre-writing, (ii) during mentoring conversations to explore ideas, and (iii) during the sharing of written drafts with the supervisor and supervisory committee. Mentoring and scaffolding existed in the form of timely and tailored feedback on writing with a focus on incremental and continual improvement (Carless et al., 2024; Kumar and Coe, 2017; Larcombe et al., 2007; Polkinghorne et al., 2023). As Shawn stated, "the whole idea was that you are supposed to grow. It was progress," while Jesse indicated, "I felt like it was okay to send something in that wasn't perfect." Participants indicated that regularly scheduled and frequent online meetings with their supervisor regarding their writing led to the continual identification of next steps including the intentional preparation for goals like candidacy and passing the final oral exam. The "meetings" happened in a variety of channels, depending on the need, by telephone, text, on Zoom, or working synchronously and asynchronously in shared Google docs leading to a "feedforward" approach where "student improvement was a key goal." Participants placed the supervisor at the center, attending to the provision of timely assessment and student application of feedback in service of the long-term goal (Sadler et al., 2022, p. 9). Participants also highlighted the importance of feedback that they saw as non-judgemental and that came in the form of questions, encouragement, suggested resources, and a focus on process. Leslie confirmed, "the feedback was there, and encouragement and resources [Supervisor] would send my way." There was also a recognition that sometimes scaffolding could involve direct instruction and was linked to assisting students in making sense of methodology, data collection, and analysis possibilities, and directing students where to go next in their writing. Alex described their experience with direct instruction: "It felt better having someone say to me, that's what you are doing. That's *not* what you are doing. Focus on this." This combination of iterative cycles, timely feedback, and diverse modes of online support helped students to progress steadily through their academic writing, reinforcing both their confidence and their scholarly development.

Graduates acknowledged the importance of the supervisors' provision of *timely* feedback (Carter et al., 2020; Can and Walker, 2014; Kar, 2024; Lim et al., 2019; Naidoo et al., 2023). As Morgan explained, "The timeliness of it was really helpful because you would get some momentum going in an aspect of your research or writing

up and having that timeliness of the feedback just really helped me when I needed to keep the momentum going.” Timely feedback not only helped students maintain momentum but also reinforced their confidence in navigating the complex stages of their research and writing.

In contrast, the lack of timeliness in supervisor feedback was also noted. Reflecting on an adverse experience of a colleague in comparison to their own, Leslie stated:

Like he could wait months before he got feedback. That's not feedback. That's nothing. Because by now you've lost your train of thought. Not just your train of thought, but you've found other research, more literature. Now you're bringing that in. Now you're going in a different direction.

Delayed feedback can have significant consequences. It can cause doctoral students to experience confusion and misalignment in the research process, which can hinder their progress and negatively impact the overall direction of the dissertation.

Participants described how their supervisor's knowledge of the field was a key component to their steady progress in writing. One praised their supervisor's mentorship as vital in promoting their emerging coherence of thought and expression. Jesse stated their supervisor had “such a depth of knowledge in terms of, you know how this fits to this, and what about this, and consider this angle.” The supervisor's expertise not only helped refine students' ideas but also facilitated a deeper understanding of the interconnectedness of concepts, which enhanced the overall quality of the student's academic writing.

One of the five participants spoke about the *desire* for such a relationship with their supervisor; however, their lived experience suggested something different. Alex recounted their supervisor's focus on procedural errors in their writing, stating, “my advisor is really good at, like, catching me on the APA and stuff but in terms of the content, that was a real struggle for me.” Alex articulated a supervisory practice that portrayed writing as a focus on mechanics rather than the development and structuring of ideas (Darland et al., 2024; Ondrusek, 2012). In addition, doctoral students not only often face challenges in the writing process, they also must contend with conceptual and methodological challenges, as well as an array of feelings associated with being a novice scholar (Calle-Arango and Ávila Reyes, 2023; Colombo, 2018; Gimenez et al., 2024).

Practice 3: building on coursework and program structures as springboards for academic writing

A third practice participants identified as supportive was the instruction and practice with academic writing through specific coursework and program structures. Participants found it helpful when program structures (Friesen and Jacobsen, 2021) linked coursework with dissertation writing. Establishing a collaboratory of practice fostered supportive feedback networks and provided shared experiences. Opportunities to engage in research as a research assistant also enhanced the academic writing process. The academic preparation gained through practice writing in coursework was a vital part of student growth in that it provided not only procedural support,

for example with APA and ethics preparation, but it offered early writing experiences as a learning and practice opportunity (Stevens and Caskey, 2023). Shawn described how “the research courses were really my reflection courses which surprised me. I thought they would be other spaces because they became very personal. It was about who are you as a researcher and which methodological approach are you going to take?” Shawn's personal reflection through coursework enabled students to gain clarity on their research identity and methodological choices, ultimately helping them progress with their dissertation writing.

Leslie articulated how linking the coursework with aspects of dissertation writing supported their learning, even though they did not utilize the writing from the course work per se in their dissertation: “It helped me to learn how to do it, but I had to totally redo that [in the dissertation]. So, was the content what it is now? No. Was the process? Yes. I learned the process.” By linking aspects of coursework with writing the dissertation, the participants described connections and structured opportunities that promoted thinking and garnered feedback on their initial ideas. Morgan indicated:

One of the major assignments was to start fleshing out your research plan. And then we presented this to the class and received feedback from them. So, it was a way to get peer support and kind of unpacking and thinking [about] our ideas.

These experiences helped to scaffold students' transition from structured coursework to independent dissertation writing, ensuring they were better equipped to refine their research approach.

Variability in the quality of instruction and course design, however, also presented challenges. Leslie stated, “Were there some gaps? Yes. And that would be dependent again on who is leading which course and how it is set up.” Shawn also described the challenge associated with varying quality in instruction and course experiences.

It was course design. It was the way it was taught. We didn't even receive feedback on our projects, like there were many issues. [It] still makes me uncomfortable. And we still talk about it as a cohort because it really did affect our choices for our research because we didn't understand what to do.

Shawn articulated, however, how their supervisor responded to the challenge. Based on a negative course experience, Shawn's supervisor advised them to take an additional course to address their learning gap. These inconsistencies in course design and instructional quality had a direct impact on students' ability to effectively plan and carry out their research, underscoring the importance of well-structured and supportive course designs and academic environments.

For Alex, who experienced coursework as a positive structure that supported their thinking and learning, they also indicated these experiences were disconnected from writing the dissertation. “After the 2 years of coursework, we started writing our dissertation and it was us and our advisor. And so, all of that was great up until it was time to write our dissertation.” In this case, Alex viewed coursework and dissertation writing as separate processes, and it was the latter where the supervisory relationship came into prominence and was not always experienced as positive or helpful for their progress.

Leadership in program design that explicitly links coursework and program structures to enhance opportunities for doctoral students to

develop self-confidence and become more familiar with academic writing conventions (Friesen and Jacobsen, 2021), while developing a base of discipline and methodological knowledge (Stevens and Caskey, 2023), is key to providing an academic environment and research culture that supports effective supervisory practice.

Practice 4: providing diverse models of academic writing for various purposes

A fourth practice identified by participants as important was providing access to diverse models and examples of writing that students would experience in academia (Carter et al., 2020; Eaton and Dombroski, 2022; Sarnecka et al., 2022). Supervisors, committee members, and other faculty modeled writing for various purposes, from proposal writing to grant writing to award and fellowship applications, to the academic writing process involved in preparing individual papers for peer review. Shawn described the importance of support when they applied for and received, with their supervisor's guidance, an international fellowship related to their doctorate. "It was a way of modeling what the expectation was in academia, because coming from [a workplace setting], I had no idea what the expectations were." Instead of starting from scratch, the doctoral student was scaffolded into this form of academic writing for a particular purpose (fellowship application) using examples provided by the supervisor from within the discipline.

Jesse augmented their own research and writing by building on their supervisor's work within a national professional organization, which provided a springboard and inspiration for doctoral study. "I think I really lucked out working with [my supervisor] and some of the work that they'd done with [a National Organization]." Jesse acknowledged how the writing experience assisted them in other aspects of their professional work following completion of their degree. "But even just developing that confidence with a research proposal helped me to write some other grants after, right?" These experiences not only supported Jesse's immediate academic success but also had long-term professional benefits, reinforcing the value of strong mentorship and applied learning in academic settings.

In the absence of mentoring from their supervisor, however, Alex sought writing models on their own. Alex described how they "literally lived at the library for two years." Asked to explain in more detail, they stated, "I downloaded and read several dissertations to kind of help me because I did not know. Like you really do not know." They also recounted how later in the doctoral program they used a colleague's dissertation as a prototype. "It was her dissertation that I used as a template for, what should it sound like, how should the layout of it be?" Given the lack of mentorship from their supervisor, Alex was left to search out and analyze writing models on their own, resulting in an extended time in the program, and their constant questioning of whether they could complete a dissertation at all.

The expectations and "environment of doctoral study has become more writing-rich than ever before" (Aitchison and Paré, 2012, p. 13; Burford et al., 2021). Growing pressure to produce a variety of quality academic texts associated with and beyond the dissertation range from conference proposals to ethics and grant applications, to journal articles, and is an assumption of modern doctoral participation. For these doctoral graduates, the models provided to them, or accessed by them, offered opportunities to build competence and confidence not

only in terms of the structures and syntax of academic writing, but also in relation to the mores and traditions of academic deportment.

Research question 2: what emotions were associated with effective online doctoral student academic writing relationships?

Several participants reported positively on the mentoring practices, while a few lamented on the absence of the mentoring practices. Next, we present a synthesis of emotions expressed by doctoral graduates when describing experiences with online academic writing in relation to the literature. We concur with other scholars that it is important for supervisors' to acknowledge emotions as potential "affective barriers" to writing (Goodson, 2023; Lim et al., 2019; Ondrusek, 2012, p. 182). In our analysis, however, we found that positive emotions were most common in the data. We posit this finding was due to an interview focus on supervisory practices that supported doctoral students. For example, of the ten most frequently cited emotions, the majority were positive. Our analysis suggests that supervisory practices that focus on fostering trust and confidence can play a critical role in mitigating the negative emotions often associated with academic writing.

Overall, we found the most prevalent emotions revealed in participant comments were positive, with a deep sense of gratitude and appreciation for their supervisor. When reflecting on the overall experience, Shawn stated, "I had a wonderful experience ... I always felt it was a team." Jesse suggested their supervisor was the key to their completion: "But if I did not have that person that I trusted, it would have been hard to get through." Leslie described their supervisor as "so dedicated, crazy amazing" and their experience as, "A good life changing [which] has left me with, you know, the yearning for more." The importance of relational trust, intentionally established and fostered in the online environment, as a key component not only in academic writing success, but successful completion of the doctoral degree, emerged strongly in the data.

Most doctoral graduates described how their supervisor listened to their ideas, took time to understand their needs, and provided personalized and timely support with their academic writing. Doctoral graduates expressed plenty of trust in their supervisor: trust they would respond in a timely manner; trust the feedback provided would be given in the spirit of improvement, revision (seeing again), and refinement (making more clear), and trust the feedback was based on a deep understanding of the field (because as Shawn stated, "you do not know what you do not know"). When that trust was lacking, (as in the case of Alex) there was a very real feeling of isolation and abandonment.

In essence, most of the doctoral graduates expressed appreciation for supervisors who extended the very best of themselves to propel the writing forward, who supported the ideas that were just emerging, and who were deeply invested in the students becoming proficient writers.

Although mostly positive, emotions such as frustration, isolation, and vulnerability, did present themselves during the interviews and often were described in visceral terms. For example, Shawn mentioned their fear and frustration at a lack of understanding of data analysis procedures that blocked their writing.

I wasn't able to write until I got through it [data analysis]. And I think that was really scary. Because then it was that moment when you think, am I ever going to get this done? Because I don't

know how. But the writing didn't happen or the [writing] tidal wave didn't happen until after [they received support with data analysis].

This same challenge was described, but with more emotional intensity, by Alex.

But I almost feel traumatized because you're, so at least in our cohort, you're so out there on your own in terms of collecting the data and sifting through it all and getting through it all. Like it's not like you have money to hire people to help you or you're sitting there talking it out with someone and you're in it together. Like you are kind of in it alone and that's kind of a really daunting position to be in.

A lack of understanding of data analysis was described by Alex, who shared the sense of isolation and overwhelm associated with navigating data collection and analysis alone. Alex's accounts underscore the emotional toll that the research process can take on students and highlight the critical need for timely support and ongoing guidance from supervisors to alleviate these pressures and foster academic progress.

As the participants expressed, learning to not only gain expertise in the discipline, critically analyze and synthesize data, but also “navigate the tensions between “knowledge telling, transformation and creation” (Gimenez and Thomas, 2015, p. 29; Gimenez et al., 2024, p. 2) amongst a myriad of conflicting emotions, often left doctoral students feeling alone and isolated. The feelings of intense vulnerability described by some participants speaks to the importance of foundational supervisory noticing in online relationships, with robust attention to frequent personal connections in various modalities to overcome isolation and geographical dispersion, coupled with writing and mentoring practices at critical times in the program (Jacobsen et al., 2021) to effectively scaffold doctoral students' learning to be scholars.

Discussion

A well-established body of research focuses on the structural or temporal supports for academic writing, such as the pedagogic “know-how” needed for success (Catterall et al., 2011; Everitt, 2022; Gimenez et al., 2024; Jones, 2018; Stevens and Caskey, 2023). We also contend that trust, community, and ongoing responsiveness to not only the academic, but also the emotional needs of the doctoral student writers are important elements leading to successful completion of their degree.

The importance of relational trust

The significance of trust as an essential component of the highly emotive process of academic writing (Jacobsen et al., 2021; Lim et al., 2019), however, cannot be understated as doctoral students “make major adjustments in how they view knowledge, learning, written expression, and themselves” (Ondrusek, 2012, p. 180). Underpinning the success for doctoral students' online academic writing was the relational trust built with their

supervisor - they trusted their supervisor would mentor and scaffold their writing experiences and leverage the program structures for their learning benefit and progress (e.g. the supervisor who advised Shawn to take an extra course). Four of the five doctoral graduates described the supervisor as a key leader in the collaborative online community of support that surrounded their academic writing experience, which included the supervisory committee, student peers, and instructors, enabling them to flourish and thrive as academic writers.

There is an intimacy that develops as supervisors observe their doctoral students becoming more confident scholars and writers in the collaborative presence of knowledgeable others. Rather than seeing the online doctoral program as a rite of passage where the apprentice is a neophyte researcher under the guidance of an expert in a hierarchical power relationship (Jacobsen et al., 2024a,b; Halse and Bansel, 2012), our study findings indicate that supervisors who nurtured doctoral student writers, even in virtual settings, led to deep feelings of gratitude and reciprocity. This finding suggests that intentional, effective relational practices with and beyond the supervisor that support doctoral student writing are connected to empowerment and flourishing versus hierarchy and power. Our findings align with Makhamreh and Kutsyuruba (2021), who emphasize that relational trust is fundamental in supervisory practices, particularly in fostering academic writing success. Their study highlights how trust-based relationships in supervision lead to greater student engagement, confidence, and the ability to navigate the complexities of doctoral writing. Our study further extends this by demonstrating that, even in online settings, relational trust not only influences academic progress but also mitigates emotional barriers, ultimately shaping the doctoral student experience and development as a confident academic writer.

The importance of a collaborative community of support

According to some doctoral graduates, the selection of committee members was a pivotal action by the supervisor. When supervisors sense integrity, a component of relational trust (Bryk and Schneider, 2002), among colleagues and possible committee members, they recognize shared and common views and perceive that the actions and support provided to the doctoral student by others will be consistent with their views. In supervision, this decision means going beyond the semi-private / private nature of student-supervisor relationships and one-on-one work with their students, to purposefully engaging and cultivating meaningful collaborations for supporting doctoral student writing with colleagues, while considering the students' positionality and vulnerability as learners.

As acknowledged by doctoral graduates in this study, the collaborative community of support plays a vital role in making visible effective supervision practices, and “recognizes that all participants in the doctoral process bring resources to and make demands on each other but defines their relationship as a cooperative endeavour of reciprocal responsibilities and obligations” (Halse and Bansel, 2012, p. 384; Roos et al., 2021). While supervisors have the primary responsibility for mentoring and guiding a student's development and progress, they also need to recognize and leverage supervision of doctoral students as part of a collaborative community of support (Jacobsen et al., 2021; Catterall et al., 2011; Chakraborty et al., 2021; Gimenez et al., 2024; Jones, 2018).

The findings from this case study reinforce the importance of integrating supervisory, peer, and committee feedback into a cohesive support system. This integrated approach fosters academic growth by providing doctoral students with multiple avenues to refine their ideas and writing. Our study builds on previous work by demonstrating that supervisory structures are most effective when they extend beyond dyadic relationships to include a network of support (Catterall et al., 2011; Gimenez et al., 2024). This aligns with research on collaborative learning environments, which suggests that shared intellectual engagement fosters greater motivation, deeper conceptual understanding, and a stronger sense of belonging within academic communities (Jacobsen et al., 2024a,b; Jones, 2018). These findings are particularly relevant in online doctoral programs, where sustained engagement with peers, committee members, and faculty can mitigate the isolation often associated with independent research. A holistic framework that recognizes both the academic and emotional dimensions of doctoral writing is critical in helping students persist and thrive in their programs.

The importance of ongoing feedback

Within the support community, feedforward is a crucial element when grappling with the questions, “Where am I going? How am I going? Where to next?” (Hattie and Timperley, 2007, p. 87; Sadler et al., 2022). This means that feedback on writing is sometimes provided as direct instruction and other times as specific recommendations of scholarly resources to promote thinking. The feedback should go beyond mechanics and editing to focus most importantly on content and structural coherence (Darland et al., 2024; Ondrusek, 2012), and draw on committee members’ broad range of methodological, discipline, and writing knowledge to encourage the promotion of academic writing as a tool for thinking.

Program structures and coursework can impact doctoral student writing success as indicated by some of our participants. The practice of instructors and supervisors engaging in tailored dialogue with doctoral students during completion of coursework to bring awareness and acknowledgement of the complexity of the academic writing process, while encouraging its use as a tool for thinking, is recommended. Developing regularly scheduled check-ins throughout the program to monitor doctoral student progress in relation to their understanding of theories, methodologies, and data collection and analysis processes were suggested by doctoral graduates as well as participation in diverse forms of academic writing and research (e.g., research assistantship).

In addition, program structures should be designed for students to maintain continued connections with their doctoral student cohort once the coursework is completed and they transition to the research program and dissertation writing stages. By integrating relational trust, a collaborative community of support, and iterative cycles of feedback, institutions can create more sustainable environments that foster both academic success and student well-being in online doctoral programs.

Conclusion

In this case study research, we identified four supervisory mentoring practices that positively impacted doctoral student online

academic writing experiences and their expressed feelings of trust and gratitude: (a) engaging students in a trusting, supportive community of practice; (b) engaging in regularly scheduled synchronous meetings with iterative cycles of mentoring and scaffolding; (c) using coursework and program structures as a springboard for writing, and; (d) providing diverse models and mentoring of academic writing. We also found that online doctoral graduates’ recognition and acknowledgement of these practices in their supervisors led to positive emotive responses, in particular gratitude, and a growing confidence in their academic writing abilities and emerging identities as scholars.

Limitations

There are several limitations to this case study that must be acknowledged. First, the findings are based on a small sample size of first language speakers from a single institution in one country, which may limit the transferability of the results. Second, as the findings are derived from doctoral graduates’ retrospective accounts, they may be influenced by agreement bias or selective memory, as all participants successfully completed their programs. Third, this study focuses on academic writing, which represents only one aspect of the doctoral experience. As such, the findings may not capture the full complexity of student-supervisor relationships or the broader doctoral program experience. Finally, while the study emphasizes online doctoral programs, we submit that the identified practices may be universal and relevant across both face-to-face and online modalities, but further research is needed to confirm this assumption.

Future research directions

Given the study’s limitations, future research could explore several key areas. Studies could examine how these mentoring practices operate across different cultural and institutional contexts to better understand their adaptability and effectiveness. Additionally, longitudinal research with a larger sample size could provide richer insights into how mentoring practices evolve over the course of the entire doctoral journey and their impact on both academic and professional outcomes. Another potential area of research is studying the intersection between supervisory practices and equity, diversity, and inclusion, particularly how these practices support students from historically minoritized groups. Finally, comparative studies between face-to-face and online doctoral programs could identify nuances in how mentoring practices are enacted in different modalities, offering a deeper understanding of their universal or context-specific nature.

Significance of the study

This study contributes to the growing body of research on fostering strong doctoral student-supervisor relationships, which is the importance of cultivating a collaborative community of support for academic writing (Jacobsen et al., 2021) and the importance in building relational trust between supervisor and doctoral student (Friesen et al., 2022). While aligning with prior studies on the critical

role of supervisors in supporting doctoral students' academic writing and successful completion of their program (Jacobsen et al., 2021; Polkinghorne et al., 2023), a key contribution of this study is the demonstrated value and importance of an online collaborative community of support for doctoral students' online academic writing (Kar, 2024; Kirkpatrick, 2019; Naidoo et al., 2023). A collaborative community of support includes regular access to and support of supervisory committee members, course instructors, doctoral student peers, and the doctoral cohort, coupled with sound program structures (Friesen and Jacobsen, 2021), to cultivate and advance doctoral student growth as academic writers. This study is significant in that it underscores the potential for joy and flourishing in doctoral education when holistic and relational approaches are employed to support students' academic journeys.

Data availability statement

The datasets presented in this article are not readily available because we wish to protect the privacy of the participants. Requests to access the datasets should be directed to sandra.becker@ucalgary.ca.

Ethics statement

The studies involving humans were approved by University of Calgary Conjoint Faculties Research Ethics Board (CFREB). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

References

- Adamek, M. E. (2015). Building scholarly writers: student perspectives on peer review in a doctoral writing seminar. *J. Teach. Soc. Work.* 35, 213–225. doi: 10.1080/08841233.2014.995333
- Aitchison, C., and Paré, A. (2012). "Writing as craft and practice in the doctoral curriculum" in Reshaping doctoral education: International approaches and pedagogies. eds. A. Lee and S. Danby (London, UK: Taylor and Francis), 12–25.
- Bastalich, W., and McCulloch, A. (2022). The ideal research degree supervisor 'can play any role': rethinking institutional orientation and induction for commencing doctoral students. *Innov. Educ. Teach. Int.* 61, 581–596. doi: 10.1080/14703297.2022.2158117
- Belcher, W. (2019). Writing your journal article in twelve weeks: A guide to academic publishing success. 2nd Edn. Chicago: The University of Chicago Press.
- Berdanier, C. G. (2021). Linking current and prospective engineering graduate students' writing attitudes with rhetorical writing patterns. *J. Eng. Educ.* 110, 207–229. doi: 10.1002/jee.20368
- Braun, V., and Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qual. Res. Sport Exerc. Health* 11, 589–597. doi: 10.1080/2159676X.2019.1628806
- Braun, V., and Clarke, V. (2020). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qual. Res. Psych.* 18, 328–352. doi: 10.1080/14780887.2020.176923
- Braun, V., Clarke, V., Hayfield, N., Davey, L., and Jenkinson, E. (2022). "Doing reflexive thematic analysis" in Supporting research in counselling and psychotherapy. eds. S. Bager-Charleson and A. McBeath (London, UK: Palgrave Macmillan), 19–38.
- Bryk, A., and Schneider, B. (2002). Trust in schools: A core resource for improvement. New York: Russell Sage Foundation.
- Burford, J., Amell, B., and Badenhorst, C. (2021). "Introduction: the case for re-imagining doctoral writing" in Re-imagining doctoral writing. eds. C. Badenhorst, B. Amell and J. Burford (Fort Collins, CO: University Press of Colorado), 3–28.
- Burkholder, G. J., and Bidjerano, M. (2023). Getting them done: a qualitative study of characteristics of mentors successful in graduating online students in a distance education context. *Mentor. Tutor.* 31, 489–511. doi: 10.1080/13611267.2023.2225393
- Calle-Arango, L., and Ávila Reyes, N. (2023). Obstacles, facilitators, and needs in doctoral writing: a systematic review. *Stud. Cont.* 45, 133–151. doi: 10.1080/0158037X.2022.2026315
- Can, G., and Walker, A. (2014). Social science doctoral students' needs and preferences for written feedback. *High. Educ.* 68, 303–318. doi: 10.1007/s10734-014-9713-5
- Carless, D., Jung, J., and Li, Y. (2024). Feedback as socialization in doctoral education: towards the enactment of authentic feedback. *Stud. High. Edu.* 49, 534–545. doi: 10.1080/03075079.2023.2242888
- Carter, S., Guerin, C., and Aitchison, C. (2020). Doctoral writing: Pleasures, practices, processes. Singapore: Springer.
- Catterall, J., Ross, P., Aitchison, C., and Burgin, S. (2011). Pedagogical approaches that facilitate writing in postgraduate research candidature in science and technology. *J. Univ. Teach. Learn. Prac.* 8, 1–12. doi: 10.5376/1.8.2.7
- Chakraborty, D., Soyoo, A., Moharami, M., Utami, A. D., Zeng, S., Cong-Lem, N., et al. (2021). Feedback as a space for academic social practice in doctoral writing groups. *Educ. Dev. Psychol.* 38, 238–248. doi: 10.1080/20590776.2021.1972764
- Cleland, J., MacLeod, A., and Ellaway, R. H. (2021). The curious case of case study research. *Med. Educ.* 55, 1131–1141. doi: 10.1111/medu.14544
- Colombo, L. (2018). The role of social relations in the making of a doctoral thesis. *IJERI* 10, 366–378.
- Consalvo, A., and Rueter, J. (2024). Supporting graduate students' writing in online courses. *J. Coll. Read. Learn.* 54, 14–31. doi: 10.1080/10790195.2023.2275051
- Darland, D. C., Gisi, E. M., Hampton, J. R., Huang, H., Kantonen, L. M., Kong, D., et al. (2024). Student-centered approaches to breaking through scientific writing barriers. *J. Coll. Sci. Teach.* 53, 445–453. doi: 10.1080/0047231X.2024.2373022

Author contributions

SB: Writing – original draft, Writing – review & editing. MJ: Writing – original draft, Writing – review & editing. SF: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Eaton, C., and Dombroski, J. (2022). Supporting graduate writers through a writing commons. *AB J. Educ. Res.* 68, 499–514. doi: 10.55016/ojs/ajer.v68i4.73428
- Everitt, J. (2022). How involved should doctoral supervisors be in the literature search and literature review writing? *Teach. High. Educ.* 28, 894–901. doi: 10.1080/13562517.2022.2145468
- Friesen, S., Becker, S., and Jacobsen, M. (2022). “Online doctoral student-supervisory relationships: exploring relational trust” in *Online learning and teaching from kindergarten to graduate school*. eds. M. Jacobsen and C. Smith (Canadian Association for Teacher Education), 282–306.
- Friesen, S., and Jacobsen, M. (2021). Collaborative design of professional graduate programs in education. *Int. J. Des. Learn.* 12, 64–76. doi: 10.14434/ijdl.v12i1.25778
- Gimenez, J., and Thomas, P. (2015). A framework for usable pedagogy: case studies towards criticality, accessibility and visibility. *Working With Academic Literacies: Case Studies Towards Transformative Practice*, eds. T. Lillis, K. Harrington, M. Lea, & S. Mitchell, Anderson, SC: ParLOUR Press, 29–44.
- Gimenez, J., Paterson, R., and Specht, D. (2024). Doctoral writing through a trajectorial lens: an exploratory study on challenges, strategies and relationships. *High. Educ.* 87, 491–508. doi: 10.1007/s10734-023-01019-7
- Goodson, P. (2023). *Becoming an academic writer: 50 exercises for paced, productive, and powerful writing*. 3rd Edn. Thousand Oaks, CA: SAGE.
- Halse, C., and Bansel, P. (2012). The learning alliance: ethics in doctoral supervision. *Oxf. Rev. Educ.* 38, 377–392. doi: 10.1080/03054985.2012.706219
- Hattie, J., and Timperley, H. (2007). The power of feedback. *Rev. Educ. Res.* 77, 81–112. doi: 10.3102/003465430298487
- Huerta, M., Goodson, P., Beigi, M., and Chlup, D. (2017). Graduate students as academic writers: writing anxiety, self-efficacy and emotional intelligence. *High. Educ. Res. Dev.* 36, 716–729. doi: 10.1080/07294360.2016.1238881
- Hutton, L., Olejnik, M., and Kunkel, M. C. (2024). Navigating contradictions while learning to write: a disciplinary case study of a first-term doctoral writer. *ATD* 21, 21–42. doi: 10.37514/ATD-J.2024.21.1.03
- Jackman, P. C., Sanderson, R., and Sisson, K. (2024). “Transition to doctoral study” in *Prioritising the mental health and wellbeing of doctoral researchers: Promoting healthy research cultures*. eds. J. Creaton and O. Gower (New York, NY: Routledge), 32–50.
- Jacobsen, M., Friesen, S., and Becker, S. (2021). Online supervision in a professional doctorate in education: cultivating relational trust within learning alliances. *Innov. Educ. Teach. Int.* 58, 635–646. doi: 10.1080/14703297.2021.1991425
- Jacobsen, M., Friesen, S., and Becker, S. (2024a). Learning doctoral supervision in education: a case study of on-the-job development of effective mentoring practices. *Int. J. Doc. Stud.* 19:012. doi: 10.28945/5375
- Jacobsen, M., Neutzling, N., Lorenzetti, L., Oddone Paolucci, E., Nowell, L., Clancy, T., et al. (2024b). Examining graduate student perspectives on supervision and peer mentoring across four professional faculties. *Stud. Grad. Postdoc. Educ.* doi: 10.1108/SGPE-03-2024-0035 [ahead of print].
- Jones, D. R. (2018). A proposed systems model for socializing the graduate writer. *Stud. High. Educ.* 43, 173–189. doi: 10.1080/03075079.2016.1160276
- Kar, R. K. (2024). An idea to explore: cultivating the art of proposal writing among graduate students. *Biochem. Mol. Biol. Educ.* 52, 369–372. doi: 10.1002/bmb.21822
- Kirkpatrick, K. J. (2019). Online doctoral students writing for scholarly publication. *Comput. Compos.* 52, 19–36. doi: 10.1016/j.compcom.2019.01.012
- Kumar, S., and Coe, C. (2017). Mentoring and student support in online doctoral programs. *Am. J. Dist. Educ.* 31, 128–142. doi: 10.1080/08923647.2017.1300464
- Larcombe, W., McCosker, A., and O’Loughlin, K. (2007). Supporting education PhD and DEd students to become confident academic writers: an evaluation of thesis writers’ circles. *J. Univ. Teach. and Learn. Pract.* 4, 61–71. doi: 10.53761/1.4.1.6
- Lavelle, E., and Bushrow, K. (2007). Writing approaches of graduate students. *Educ. Psychol.* 27, 807–822. doi: 10.1080/01443410701366001
- Lee, K., Zawacki-Richter, O., and Cefa Sari, B. (2024). A systematic literature review on technology in online doctoral education. *Stud. Cont. Educ.* 46, 38–64. doi: 10.1080/0158037X.2022.2135499
- Lim, J., Covrig, D., Freed, S., De Oliveira, B., Ongo, M., and Newman, I. (2019). Strategies to assist distance doctoral students in completing their dissertations. *IRRODL* 20, 192–210. doi: 10.19173/irrodl.v20i5.4532
- Makhamreh, M. A., and Kutsyuruba, B. (2021). The role of trust in doctoral student-supervisor relationships in Canadian universities: the students’ lived experiences and perspectives. *J. High. Educ. Theory Pract.* 21, 124–138.
- Maldonado, L. G., Dolfi, J. J., Bartlett, J. E., and Bartlett, M. E. (2021). Forward momentum: providing supportive space for EdD students’ dissertation progression through weekly online writing sessions. *Impact. Educ.* 6, 13–20. doi: 10.5195/ie.2021.186
- Merriam, S., and Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation*. 4th Edn. Hoboken, NJ: Jossey Bass.
- Miller, E. M., Porter, J. E., and Barbagallo, M. S. (2023). Simplifying qualitative case study research methodology: a step-by-step guide using a palliative care example. *Qual. Rep.* 28, 2363–2379. doi: 10.46743/2160-3715/2023.6478
- Naidoo, K., Quaynor, L., and Shen, Y. (2023). Doctoral students’ experiences, self-efficacy, and sense of belonging related to academic writing in an online program. *Educ. Sci.* 13, 1–22. doi: 10.3390/educsci13111097
- Ondrusek, A. L. (2012). What the research reveals about graduate students’ writing skills: a literature review. *J. Educ. Lib. Info. Sci.* 53, 176–188.
- Oxford English Dictionary. (2024) Trust. Available at: https://www.oed.com/dictionary/trust_n?tab=meaning_and_use#17393047 (Accessed 12 September, 2024).
- Polkinghorne, M., Taylor, J., Knight, F., and Stewart, N. (2023). Doctoral supervision: a best practice review. *Encyclopedia* 3, 46–59. doi: 10.3390/encyclopedia3010004
- Roos, L., Löfström, E., and Remmik, M. (2021). Individual and structural challenges in doctoral education: an ethical perspective. *Int. J. Doc. Stud.* 16, 211–236. doi: 10.28945/4738
- Sadler, I., Reimann, N., and Sambell, K. (2022). Feedforward practices: a systematic review of the literature. *Assess. Eval. High. Educ.* 48, 305–320. doi: 10.1080/02602938.2022.2073434
- Saldaña, J. (2021). *The coding manual for qualitative researchers*. Thousand Oaks, CA: SAGE.
- Sarnecka, B. W., Silva, P. N., Coon, J., Vickers, D. C., Goldstein, R. B., and Roudier, J. N. (2022). Doctoral writing workshops: a pre-registered, randomized controlled trial. *Innov. High. Educ.* 47, 155–174. doi: 10.1007/s10755-021-09574-6
- Simpson, S. (2012). The problem of graduate-level writing support: building a cross-campus graduate writing initiative. *WPA* 36, 95–118.
- Stevens, D., and Caskey, M. M. (2023). Building a foundation for a successful doctoral student journey: a scholarship of teaching and learning investigation. *Innov. High. Educ.* 48, 433–455. doi: 10.1007/s10755-022-09624-7
- Wei, J., Carter, S., and Laurs, D. (2019). Handling the loss of innocence: first-time exchange of writing and feedback in doctoral supervision. *High. Educ. Res. Dev.* 38, 157–169. doi: 10.1080/07294360.2018.1541074
- Wikeley, F., and Muschamp, Y. (2004). Pedagogical implications of working with doctoral students at a distance. *Distance Educ.* 25, 125–142. doi: 10.1080/0158791042000212495
- Winberg, C., Dippenaar, H., Engel-Hills, P., and Phillips, H. (2023). Writing together alone: digitally connected ‘snack writing’ for progressing academic writing. *Read. Write. J. Read. Assoc. SA* 14:414. doi: 10.4102/rw.v14i1.414
- Yin, R. K. (2018). *Case study research: Design and methods*. 6th Edn. Thousand Oaks, CA: SAGE.
- Zimmer, W. K., Chang, C. N., Semma, B. M., and Fowler, D. (2022). Developing graduate writing habits and skills: establishing writing sessions with STEM graduate students. *Coll. Teach.* 70, 133–144. doi: 10.1080/87567555.2021.1909524



OPEN ACCESS

EDITED BY

Rui Alexandre Alves,
University of Porto, Portugal

REVIEWED BY

E. Michael Nussbaum,
University of Nevada, United States
Xiangqing Luo,
Jiangsu University of Technology, China

*CORRESPONDENCE

Anisha Singh
✉ asingh56@sfsu.edu

RECEIVED 13 January 2025

ACCEPTED 16 April 2025

PUBLISHED 15 May 2025

CITATION

Singh A, Sun Y, Alexander PA and Zhao H
(2025) Componential modeling of
argumentative essay writing from multiple
online sources: a Bayesian network approach.
Front. Psychol. 16:1560088.
doi: 10.3389/fpsyg.2025.1560088

COPYRIGHT

© 2025 Singh, Sun, Alexander and Zhao. This
is an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Componential modeling of argumentative essay writing from multiple online sources: a Bayesian network approach

Anisha Singh^{1*}, Yuting Sun², Patricia A. Alexander³ and
Hongyang Zhao⁴

¹Department of Psychology, San Francisco State University, San Francisco, CA, United States,

²Department of Teaching, Learning & Curriculum, University of North Florida, Jacksonville, FL,

United States, ³Department of Human Development and Quantitative Methodology, University of

Maryland, College Park, College Park, MD, United States, ⁴Department of Educational Psychology,

University of Illinois at Urbana-Champaign, Champaign, IL, United States

Introduction: Writing argumentative essays using multiple sources is a critical skill for college students, yet it remains a significant challenge. Despite previous research acknowledging this difficulty, the specific dynamics of the argumentative essay writing process and where breakdowns occur remain unclear.

Methods: College students wrote argumentative essays on a controversial topic after reading multiple documents. The data were fitted to two competing theory-based Bayesian networks, a method highly suited to the modeling of cognitive processes identified with argumentative writing.

Results: The best-fitting model showed that the argumentative essay task is both initiated and sustained by higher-order integration components. This model lends support to the description of the process of argumentation writing from multiple documents put forth by the stage-based Integrated Framework of Multiple Texts. Further, we found that the process of argumentation falters due to students' inability to frame counterarguments and their non-optimal critical analysis.

Discussion: This research not only enriches our understanding of the mechanics of argumentative writing from multiple sources, but the innovative Bayesian approach could lead to further refinement of the model by future researchers.

KEYWORDS

argumentation, multiple source use, argumentative essay writing, Bayesian network analysis, multiple documents, college students

Introduction

The importance of oral or written argumentation is well-established in the literature. Educational research has demonstrated that the ability to formulate cogent arguments is critical to learning across domains (Asterhan and Schwarz, 2007; De La Paz, 2005; Wiley and Voss, 1999). Further, with its emphasis on evidence and consideration of varying and contradictory perspectives, argumentation is at the heart of a democratic education (Gutmann, 1999; Hess and Avery, 2008). It should come as no surprise, therefore, that this manner of thinking and reasoning has been the subject of philosophical and psychological analysis since the time of Aristotle. In his most famous work on the topic of oral argumentation written in the 4th century BCE, *The Art of Rhetoric*, Aristotle drew on the history of ancient logic and intricately analyzed the art of persuasion. *The Art of Rhetoric* is the foundational treatise on which modern argumentation theory is based (Aristotle, 4th century BCE, 2019; Perelman and Olbrechts-Tyteca, 1969; van Eemeren, 2013).

The importance of argumentation is further evidenced in the attention it is paid in educational policies and practices intended to promote learner development and the overall wellbeing of society (Jiménez-Aleixandre and Erduran, 2007; Asterhan and Schwarz, 2016). For example, the K–12 Common Core State Standards for writing lists the ability to “write arguments to support claims in an analysis of substantive topics or texts using valid reasoning and relevant and sufficient evidence” as a curricular goal (National Governors Association Center for Best Practices and Council of Chief State School Officers, 2010, p. 18). Despite these standards and policy mandates, far too many students struggle with argumentation, as numerous research studies have documented (e.g., Kuhn, 1991; McCann, 1989; Means and Voss, 1996). Researchers have demonstrated that students are unable to recognize and apply argumentative text structures (Chambliss and Murphy, 2002; Freedman and Pringle, 1984); have difficulty providing appropriate evidence to justify claims or positions (Kuhn and Modrek, 2021; List et al., 2022); and fail to offer counterarguments or rebuttals (Ferretti and Fan, 2016; Leitão, 2003; Mason and Scirica, 2006).

Such academic challenges are amplified when argumentation takes the form of a written product (National Assessment of Educational Progress, 1994, 1999, 2002). This is because skilled writing is in itself a complex activity depending on cognitive, contextual, and motivational factors (Bereiter and Scardamalia, 1987; Galbraith and Torrance, 1998; Graham, 2018). Thus, producing a solid piece of argumentation in writing entails the transformation of knowledge and requires skills underlying composition, cognizance of the constraints imposed by the argumentative genre, specific topic, and the audience being addressed, while being driven and efficacious at managing the challenges of the writing task.

An additional source of difficulty in constructing arguments arises from the use of multiple texts, particularly with online resources (Alexander and the Disciplined Reading and Learning Research Laboratory, 2012; Stadler, 2017). Those challenges pertain to the proliferation of information available online, multiple perspectives on the same issue, and varying degrees of source credibility and content accuracy (Braasch et al., 2014; McGrew, 2021). Consequently, students called upon to craft argumentative essays from multiple documents must be able to evaluate and integrate the information from multiple documents, even before they begin to write. Cognizant of these challenges, researchers have turned their attention on examining and promoting competencies related to using multiple sources. These competencies include the ability to assess the reliability of sources and establish connections among ideas across different documents (Anmarkrud et al., 2013; Braasch and Bråten, 2017; Britt and Rouet, 2012). Developing these skills is crucial for crafting argumentative essays within the pluralistic information landscape of the internet.

The current study builds upon and extends the aforementioned research on argumentation, particularly in its written form within the context of using multiple sources. It employs an innovative statistical method—Bayesian Network analysis—to model the componential processes involved in producing a quality argumentative essay from multiple documents.

Specifically, in this study, students were required to access, read, and integrate information from a library of online documents that varied in both source and content credibility. We selected only original documents from the internet to create the library. We carefully curated documents that represented various combinations of source and content credibility. For instance, we included documents from credible sources that were found to contain content of questionable credibility. This is different from previous studies where only the credibility of sources, and not the content of those sources, was manipulated (e.g., Ecker and Antonio, 2021; Sparks and Rapp, 2011; van Boekel et al., 2017). Further, we employed Bayesian Network analysis, a probability-based technique, that allows for the modeling of causal relations among components and make predictions about the relative importance of each component to the production of a quality argumentative essay. The rationale for the use of this more novel technique was to shed light on the complex interrelations among the components that constitute the argumentative writing process. This analysis would also allow us to identify components of argumentation that seem particularly challenging for students.

Due to the fact that we used Bayesian Network analysis as a theoretically driven approach where key components were specified prior to modeling, we first discuss argumentation and the components entailed in its execution. We then describe the particular framework of multiple source use into which the writing of an argumentative essay was embedded. Finally, given the somewhat novel modeling procedure we apply, we briefly overview Bayesian Network analysis.

The process of argumentation

Argumentation is a complex process that has been studied across multiple disciplines, each offering unique perspectives and models. While philosophical approaches often emphasize logic, resulting in the well-known inductive and deductive argument structures, this study adopts a broader, dialectical view of argumentation (van Eemeren and Grootendorst, 2004). This choice is motivated by our focus on everyday contexts involving controversial social topics, rather than purely scientific or philosophical debates.

In the dialectical approach, van Eemeren et al. (1996, p. 5) define an argument as “a verbal and social activity of reason aimed at increasing (or decreasing) the acceptability of a controversial standpoint for a listener or reader, by putting forward a constellation of propositions intended to justify (or refute) the standpoint before a ‘rational judge’.” This definition emphasizes argumentation as a communicative, rational activity aimed at influencing standpoints through justification and refutation of anticipated counterarguments.

The dialectical view frames argumentation as a goal-directed, interactional process where two or more parties engage to resolve a conflict of opinion. An argument comprises a claim that is supported by evidence, anticipates potential challenges, and is strengthened by addressing counterarguments (Walton, 2007).

Given the cognitive focus of our study, we aim to elucidate the process of argumentation from a cognitive perspective, rather

than examining the textual structure of arguments produced by our students. We posit that the generation of claims, the provision of justifications, and the formulation of counterarguments are cognitive processes that manifest as observable features in the resulting text. This psychological approach conceptualizes argumentation not merely as a social and verbal activity, but critically as a series of complex mental operations.

Our framework posits that the cognitive processes underlying argumentation—specifically, claim formulation, evidence evaluation, and anticipation of opposing viewpoints—are reflected in both the structural and content-based elements of the argumentative text. We hypothesize that these cognitive operations leave discernible traces in the textual output, providing a window into the mental processes of the arguer (Galbraith, 1998; Van Wijk, 1998).

This cognitive-centric model allows us to investigate the intricate interplay between internal cognitive mechanisms and their external manifestations in argumentative discourse. By focusing on these cognitive underpinnings, we aim to develop a more nuanced understanding of the mental operations that drive effective argumentation, potentially offering insights into cognitive strategies that can enhance argumentative skills.

Integrated framework of multiple texts

The process of argument construction based on multiple texts has been a focus in the body of literature on multiple source use (Barzilai et al., 2021; De La Paz and Felton, 2010; Vandermeulen et al., 2020). In multiple source use (MSU) tasks that culminate in the production of argumentation, students need to read multiple texts on the focal issue and integrate information and perspectives from different documents to build their own arguments. The complex processes of using multiple texts to produce a desired outcome have been characterized by several theoretical frameworks (e.g., Documents Model Framework, Perfetti et al., 1999; Internet Information Problem-Solving model, Brand-Gruwel et al., 2009; Multiple Documents-Task-based Relevance Assessment and Content Extraction model; Rouet, 2006). Common to these models is the emphasis on how learners consolidate information within and across multiple documents to create an integrated representation of the texts and the topic. Among the various models, we turned to the Integrated Framework of Multiple Texts (IF-MT, List and Alexander, 2019) as a guide for our effort to understand the process of integrating information from multiple texts in producing quality argumentation.

The IF-MT is a comprehensive framework that was a consolidation of other existing models. The framework delineates three stages that explain the complex process underpinning students' multiple source use—preparation, execution, and production. The unfolding of argumentative writing in the MSU contexts in IF-MT's three stages is visually depicted in Figure 1.

In the first stage, *preparation*, students orient themselves by conducting task analysis to determine the requirements of the assignment at hand and begin mentally mapping the steps toward completion. Students' analysis of the task is influenced by the interplay of their individual characteristics (e.g., knowledge, interest, attitude) and external task demands. For

example, students' prior knowledge about what argumentation constitutes can influence the quality of their argumentative essays (Nussbaum, 2011).

The second stage, *execution*, is where students implement the steps for completing the task, and the stances they adopted begin to manifest in external actions. During this stage, students search, navigate, select, and read sources and then forge associations within and across documents. Through these complex processes, students engaged in an argumentative task may develop mental representations of the informational terrain of the topic by integrating diverse perspectives and supporting justifications from various documents, which may shape or reshape their stances on the topic.

The final phase of the IF-MT is the production stage, wherein students engage in the complex process of text generation. In the context of our study, this culminated in the creation of argumentative essays. Crucially, the mental representations formed during the preceding planning and execution stages significantly influence the production process. We chose to focus on this stage where text production happens because it provides a window into cognitive processes (Galbraith and Torrance, 1998).

Unlike traditional writing models (Hayes and Flower, 1980), the IF-MT posits that production is not merely a mechanical conversion of thought to text, but rather a dynamic, on-line process that plays a pivotal role in shaping the final product. This aligns with the text-production perspective, which conceptualizes the act of writing itself as a process that engenders new understanding and facilitates “knowledge transformation” (Bereiter and Scardamalia, 1987; Galbraith, 1998).

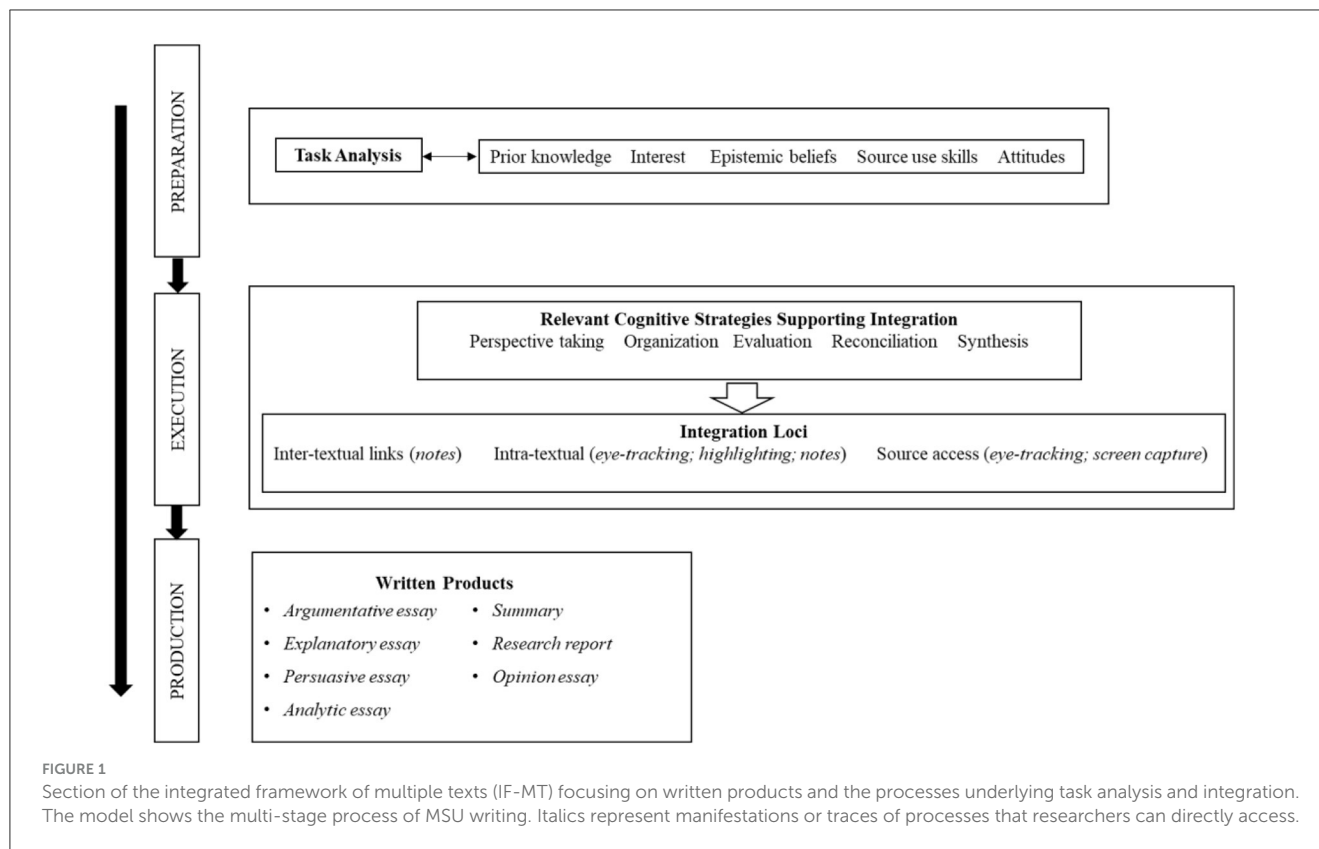
This theoretical stance suggests that the production stage offers a unique opportunity to assess the efficacy of students' comprehension of task parameters and their level of information integration. As Galbraith (1998) argues, analyzing the text as a window into cognitive processes can be a particularly fruitful approach, as it centers the intricate processes involved in text production.

In educational research and practice, written products—particularly argumentative essays—are frequently employed in multiple source use tasks (Luna et al., 2022; Mateos et al., 2018). Recognizing this, our study focused on the production stage, leveraging the IF-MT to elucidate the cognitive components that contribute to the composition of high-quality argumentative essays derived from multiple documents.

This approach allows us to examine the cognitive processes involved in text production through an analysis of their traces in the final product. By doing so, we aim to contribute to a more nuanced understanding of the complex interplay between task and integration components that underpin the creation of an argumentative essay for the broader purposes of influencing writing instruction and support.

Identifying core components of multiple-text-based argumentation

Based on the extensive literature on argumentation and multiple source use just overviewed, we identified core components of argumentative essay writing involving multiple documents.



The components included both essential elements for completing the argumentative essay task (i.e., task parameters) and core processes for achieving integration of multiple sources (i.e., integration components). Each of the task parameters and the integration components captures complex cognitive processes that are manifested in and inferred from the written product.

Task parameters

Drawing on the pragma-dialectical approach to argumentation (van Eemeren and Grootendorst, 2004), we delineated the following task parameters: (1) presenting a *claim*, (2) providing *justifications* for the claim, (3) addressing *counterarguments* through rebuttal or refutation, (4) using multiple sources. The first three parameters (claim, justifications, and counterarguments) directly reflect the dialectical nature of argumentation. The fourth parameter, using multiple sources, was deemed necessary given the multiple-text-based nature of the argumentative essay task. This parameter adds an additional layer of complexity to the task, requiring students to navigate and select multiple documents.

Further, foundational to generating a written product was students' *writing ability*. In effect, the ability to communicate through writing was judged as foundational to the production of an argumentative essay.

These delineated task components emerge from the cognitive processes underlying writer's representation of the rhetorical problem (Galbraith, 1998; Flower and Hayes, 1980). This cognitive

perspective allowed us to conceptualize the task parameters as manifestations of the mental operations involved in writing an argumentative essay.

Integration components

Integration, as defined by Alexander and the Disciplined Reading and Learning Research Laboratory (2020) is "the meaningful consolidation of elements found within and across information sources that results from the *analysis* and *synthesis* of their contents" (p. 408). This definition highlights the foundational roles of analysis and synthesis to integration. These processes can occur throughout the three stages described in the IF-MT (List and Alexander, 2019).

At the preparation stage, students may engage in preliminary task analysis as they inspect task requirements, available materials, and contextual characteristics vis-à-vis their knowledge, beliefs, and motivations. This initial assessment results in the adoption of a default stance toward task completion (e.g., critical analytic; List and Alexander, 2017). The chosen stance influences enactment of the task parameters that are reflected in the final written product.

In execution, analysis can occur when students critically evaluate the quality of the sources and their contents and identify the relations between pieces of information within and across documents. As students process the texts, they may synthesize the contents across texts depending on the consistent or conflicting nature of the information being synthesized. The depth and quality

of this analysis and synthesis manifest in the sophistication of the argumentative essay.

The production stage is where earlier cognitive processes along with the cognitive processes associated with writing become externalized in the written essays. Effective integration in writing requires coherent expression of ideas, with content-based connections between sentences and paragraphs. As Alexander and the Disciplined Reading and Learning Research Laboratory (2020) explain, “cohesion requires not only the synthesis of content across information sources, but also the analytical ability to produce inferences that bridge the informational gaps that will inevitably exist” (p. 411).

This conceptualization of cohesion adopts a cognitive perspective, aligning with the mental processes fundamental to problem-solving. Creating a coherent text necessitates that the writer maintains a goal-directed approach throughout composition. This cognitive lens emphasizes that coherence is not merely a textual feature, but rather the result of deliberate mental operations initiated in the preparation and execution stages, and externalized in the production stage.

Given the importance of these cognitive processes, we focused on three core integration components for producing an integrated written essay: (1) *critical analysis*, (2) *synthesis*, (3) *content-based overall cohesion*.

Unpacking the interplay between task parameters and integration components

Together, the enactment of the task parameters and integration components undergird the production of argumentative essays based on multiple sources. What needs to be further explored is how these components work together in the production of the written essay. We claim that the process of composing an argumentative essay unfolds in a way that certain components are prerequisites for the manifestation of other specified components. For example, in an argumentative essay, a claim must be forwarded before providing justifications or addressing potential counterarguments. The directionality of this particular process is straightforward—justification follows a claim—but the association between some of the other components is less established. For example, is critical analysis a prerequisite for synthesis or vice versa? Does critical analysis come into play before a student forms a counterargument? Therefore, in this study, we explored the specific linkages among the identified task and integration components using Bayesian network analysis. Before describing the models we tested, we present a brief description of the Bayesian networks used to formulate those models and make inferences about the processes entailed in argumentative essay writing.

Bayesian network analysis

Bayesian network analysis is a powerful statistical tool that allows for the modeling of complex causal relations among variables. At its core, a Bayesian network is a graphical model

that represents probabilistic relationships among a set of variables (Jensen, 1996; Jensen and Nielsen, 2007). It consists of two key components: structure and strength. The structure is represented by a Directed Acyclic Graph (DAG), where variables are depicted as nodes and the dependencies between them as arrows (Pearl and Russell, 2003; Murphy, 1998). The strength of these relations is quantified by conditional probability distributions representing how strongly variables in the network influences one another.

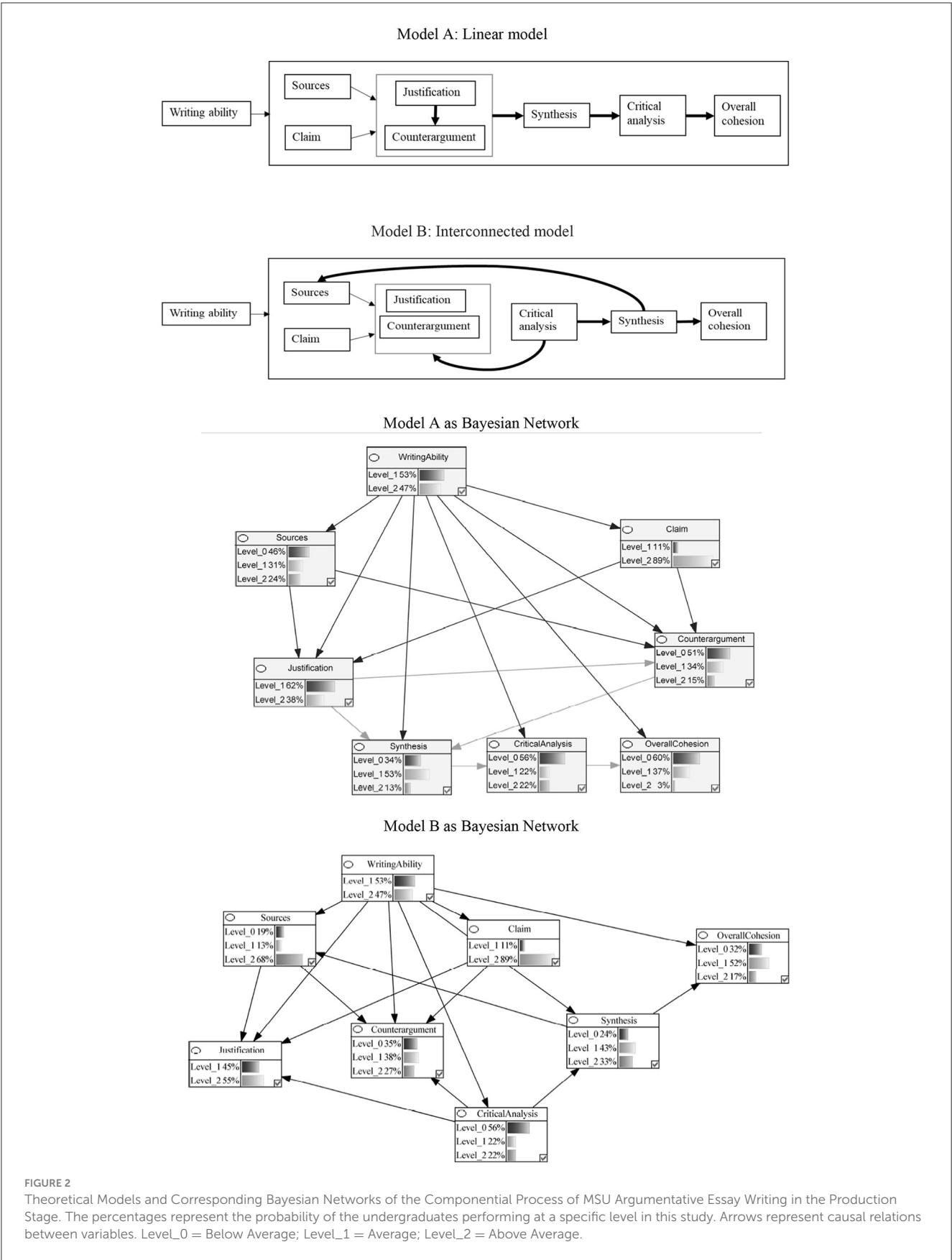
In our study of argumentative writing, we employed a hybrid approach to Bayesian network analysis. We specified the network structure *a priori*, while the parameters (strengths of relationships) were learned from data, integrating theory with computational learning. This approach of combining expert knowledge and machine learning is particularly suitable for modeling complex cognitive processes.

Our models represented various components of argumentative writing as nodes in the network, with arrows indicating the hypothesized causal relationships between these components. Figure 2 provides a visual representation of these network, illustrating how we conceptualized the process of argumentative essay writing as a series of interconnected cognitive components that leave traces in the text product.

Bayesian network analysis offers several advantages over frequentist statistical techniques. For example, a key advantage of Bayesian networks over generalized linear models is their ability to compute the impact of changes in a subset of variables that are part of an entire network or a subset of it (Pearl and Russell, 2003). In our case, we can investigate how selection of sources and synthesis of content may impact the quality of justifications, counterarguments, and overall cohesion of the essay. Bayesian network analysis also offers several advantages over traditional structural equation modeling (SEM) approaches for modeling argumentative essay writing. Specific to our study, it provides a robust method for handling ordinal data with a small sample size, while capturing complex, non-linear relations between components.

Most importantly, Bayesian networks offer a distinct advantage in result interpretation, particularly for argumentative essay writing. Unlike the continuous estimates in SEM, Bayesian networks express outcomes as probabilities of achieving specific score levels. This approach aligns more closely with how educators conceptualize student performance, making findings more intuitive and actionable. For example, a Bayesian network can directly convey the probability of a student achieving a high score on the *claim* component based on their *source* use performance. This probabilistic framework captures the nuanced relationships between components more effectively than linear estimates, better reflecting the complex nature of argumentative writing with multiple documents.

Building upon these advantages, our study leveraged Bayesian updating to further enhance our analysis (Almond et al., 2015; Pearl, 1988). Bayesian updating applies Bayes' theorem to the complex interdependencies within the network, allowing us to refine our understanding of the argumentative writing process as new data are considered. In practice, Bayesian updating involves computing the posterior probability of an event given its prior probability and likelihood function. With our model structure



specified based on theory and parameters determined from empirical data, we used this updating process to draw nuanced inferences about argumentative essay writing.

This approach transcends simple correlation analysis, enabling us to explore probabilistic causal relationships (Pearl, 2000). When we update a node—meaning we set its value to a specific state—the Bayesian network recalculates probabilities throughout the system. This updating process reveals how new information about one variable propagates to influence both its parent nodes (variables that directly affect it) and child nodes (variables it directly affects), creating ripple effects of probability changes throughout the entire network. This dynamic probability provides a comprehensive view of the interrelationships within the argumentative writing process.

For instance, we could examine how improvements in source selection might cascade through the network, affecting justification quality and overall essay cohesion. This capability provides actionable insights for educators, highlighting which components of argumentative writing are most challenging for students and how interventions in one area might impact performance in others.

Model diagnostics

To validate our Bayesian network models, we employed cross-validation, one of several diagnostic methods available for this purpose (Sinharay, 2006). While other techniques such as item fit plots, item-test statistics, and posterior predictive model checking exist, cross-validation is particularly well-suited for assessing a model's predictive power (Sinharay, 2006; Yan et al., 2003).

Specifically, we utilized leave-one-out cross-validation, an extension of the k-fold cross-validation technique, to estimate our models' goodness of fit. This method involves training the Bayesian network n times, where n equals our sample size. In each iteration, the algorithm excludes one data point, uses the remaining data to train the model, and then predicts the excluded point. This process is repeated for all n data points, providing a robust assessment of the model's predictive accuracy across our entire dataset.

This approach allowed us to evaluate how well our models predicted students' performances in argumentative writing tasks, ensuring the reliability and generalizability of our findings. For a more detailed description of our Bayesian network analysis methodology, please refer to the [Supplementary material](#).

Research questions and hypotheses

Students' struggle with argumentative writing has been well-documented and empirically explored. However, challenges with argumentation get amplified when students function in contexts with multiple documents. Despite an abundance of theoretical models of the argumentative process, there is limited understanding of how this componential process unfolds in a multiple source task. Further, we do not know where the process of argumentative writing breaks down for most students. Therefore, in this study, we posed the following research questions:

(1) *Based on Bayesian network analysis, which of the plausible theoretical models best captures the process of writing an*

argumentative essay from multiple documents? Specifically, which model - linear or interconnected - better predicts students' performance on task and integration components?

Based on the IF-MT, we hypothesized that understanding the task requirements is a prerequisite to producing an argumentative essay. The task components we included in our model were: (a) the linguistic ability to write effectively (*writing ability*), without which the student is unable to initiate the entire process of producing an argumentative essay; (b) using *sources*; (c) stating a *claim*; (d) providing *justifications*; and (e) discussing *counterarguments*. The sequential interconnections were ascertained based on how the argumentative writing unfolds. The next set of building blocks were the three theoretically determined integration components relevant to MSU written tasks: (a) *critical analysis* of the sources and contents, (b) *synthesis* of the multiple sources and the content encountered within and across documents, and (c) *overall cohesion* of the ideas presented in the written product. According to the IF-MT, critical analysis and synthesis occur in the execution stage and are evidenced in the written outcomes in the production stage when readers make intra- and inter-textual links. Although the sequence in which the task components unfold is somewhat apparent, the precise manner in which the integration components play out in the writing process needs to be investigated.

In this study, we examined two plausible models that varied in their interconnections among components: a linear model and an interconnected model. Figure 2 presents the conceptual models depicting these plausible interrelations, which were subsequently converted into Bayesian networks for analysis.

It is crucial to note that each component in our models represents a complex cognitive process, each worthy of individual study and computationally complex to model. The underlying cognitive processes are likely distributed in nature (McClelland et al., 1986). For instance, *synthesizing* ideas relies on distributed semantic memory (Galbraith, 1998). However, our focus in this study is not on the internal workings of each component, but rather on the orchestration of the components in text production.

Our emphasis on this level of analysis stems from two key considerations. First, we aim to describe cognitive processes at a level that can lead to actionable educational implications. Second, while each component and their orchestration involve complex cognitive processes, the product of each component is traceable in the essay text. Thus, we view the text as a window into how these coarser-grained components come together in the writing process.

This approach allows us to examine the architecture of the written product as a reflection of the sequential interaction of these components. By focusing on this level of analysis, we seek to bridge the gap between complex cognitive processes and observable outcomes in argumentative writing, potentially informing educational practices and interventions.

(2) *What does Bayesian updating indicate about the relative importance of the components to the writing of the argumentative essay?*

(2a) *How do early task components, particularly writing ability, formulating a claim, and source use, influence subsequent task and integration components in the argumentative writing process?*

- (2b) *What is the impact of critical analysis on other task and integration components throughout the network?*
- (2c) *How does synthesis ability affect both preceding and subsequent components in the argumentative text production?*

Using the selected Bayesian network, we estimated the probabilities associated with the sequential and bi-directional causal interrelations between the components. Specifically, we tested a series of different performance scenarios with the Bayesian network. For example, we modeled a scenario in the Bayesian network where a student exhibited the highest level of critical analysis and observed how the other interrelated components changed. Given the interconnected nature of the Bayesian network, this was akin to asking: what level of performance does a student need on task and integration components to exhibit the highest level of critical analysis in their essay? We conducted this analysis, known as *belief updating*, to determine which components were most critical for the process of argumentative essay writing.

For argumentative essay writing, we hypothesized that writing ability would have a substantial influence on all subsequent task and integration components. We expected that formulating a clear claim early in the writing process would positively impact justifications, counterarguments, and integration components. Furthermore, given that this is a multiple source use task, we predicted that effective source use would be a key component for producing a quality essay, influencing both task components (justifications and counterarguments) and integration components.

Regarding critical analysis, we hypothesized that it would show strong effects on both task and integration components (e.g., improving source selection, enhancing synthesis and justifications). This is because critical analysis operates at multiple levels—evaluating source credibility, content verification, assessing the logical connection between claims and their supporting evidence.

For synthesis ability, we hypothesized that it will have a significant impact on justifications, counterarguments, and overall cohesion. We anticipated that strong synthesis skills will be reflected in improved integration of multiple sources in the essay, contributing to a more coherent and well-supported argumentative essay.

Concerning counterarguments, we predicted that including them will be crucial for the integration process, particularly enhancing critical analysis and synthesis. We expected that strong performance in counterarguments will positively influence essay cohesion. The pivotal role of counterarguments was anticipated due to their critical importance in dialectical argumentation, where they serve to strengthen the overall argument by addressing potential objections and alternative viewpoints (Nussbaum and Schraw, 2007; Walton, 2007).

Overall, we expected that the interplay between task components (writing ability, claim formulation, source use, justifications, and counterarguments) and integration components (critical analysis, synthesis, and cohesion) will be complex and multidirectional. We anticipated that improvements in one area will potentially influence both preceding and subsequent components in the argumentative writing process, reflecting the interconnected nature of cognitive processes involved in writing argumentative essays from multiple sources.

Method

Participants

Participants in this study were 95 undergraduate students at a large Mid-Atlantic university (57 females, 1 non-binary, 1 preferred not to say). These students were enrolled in a general education course aimed at developing their learning capabilities through discussions of relevant topics (e.g., problem-solving, transfer, reasoning, and motivation) and practical learning experiences. Students represented varied majors, including arts and humanities (e.g., English, anthropology, philosophy), social sciences (e.g., criminal justice, economics, psychology), natural sciences (e.g., mathematics, physics, biology), and applied sciences (e.g., engineering, computer science, information science). The participants included freshmen (20%), sophomores (25.3%), juniors (26.3%), and seniors (28.4%), with a mean age of 20.19 ($SD = 1.47$). Their racial backgrounds were diverse, with 42.1% self-identified as White, 20% as Asian, 13.7% as Black, 4.2% as Latino, 15.8% as multiracial, and 4.2% as other races.

The study was approved by the Institutional Review board.

Multiple source use task

Students completed the multiple source use task as an assignment in the course. The task required students to read texts in a provided digital library and write an argumentative essay regarding the claim statement: “*Students today are overly dependent on technology to the detriment of their social, physical, emotional, and academic well-being.*” Specifically, students were required to select at least four out of the ten documents in the library and use what they read to compose an argumentative essay. This MSU task is an integral component of the course, as challenges of discerning credible sources and using online information for learning were key topics in the course. Students were not provided any pre-task instruction about how to read or use multiple texts to construct argumentative essays. Rather, their performance was used as a basis for post-task discussion on the challenges of engaging in such multiple document tasks, which are relatively common for college students.

The topic for the MSU task was chosen for its perceived controversy and interestingness, as reported by students ($n = 48$) enrolled in the same course in the previous semester. Those students represented similar demographic and academic backgrounds as the participants in the current study. Among a list of ten topics, the question about students’ overdependence on technology was rated as the most controversial ($M = 57.1$, $SD = 24.6$) and most interesting ($M = 68.5$, $SD = 23.0$) on a scale of 0 to 100.

Digital library

The 10-document digital library was linked to a menu that resembled a Google search page with the title, publisher, date of publication, URL, and a blurb for each document (see Figure 3). The documents linked to the menu were screenshots of original websites from the Internet with minimal modifications (e.g.,

Search Page

An Ugly Toll of Technology: Impatience and Forgetfulness

www.nytimes.com/2010/06/07/technology/07brainside.html ▼

Jun 6, 2010 - If you answered yes to any of those questions, exposure to technology may be slowly reshaping your personality. Some experts believe excessive use of the Internet, cellphones and other technologies can cause us to become more impatient, impulsive, forgetful and even more narcissistic. "More and more ...

Teenagers are better behaved and less hedonistic nowadays - The ...

<https://www.economist.com/.../21734365-they-are-also-lonelier-and-more-isolated-tee...> ▼

Jan 10, 2018 - AT THE gates of Santa Monica College, in Los Angeles, a young man with a skateboard is hanging out near a group of people who are smoking marijuana in view of the campus police. His head is clouded, too—but with worry, not weed. He frets about his student loans and the difficulty of finding a job, ...

Secrecy Is Dead. Here's What Happens Next. | WIRED

<https://www.wired.com/story/secrecy-is-dead-heres-what-happens-next/> ▼

Dec 16, 2017 - Since its founding, WikiLeaks had portrayed itself as the ultimate fourth estate—a digital drop-box where secrets could be deposited and released as public information. But in the runup to the presidential election, WikiLeaks' dispatches began to show a partisan slant. There was an email trove from a hack ...

Teens' online friendships just as meaningful as face-to ... - ScienceDaily

<https://www.sciencedaily.com/releases/2017/09/170927105416.htm> ▼

Sep 27, 2017 - Many parents worry about how much time teenagers spend texting, sharing selfies and engaging in other online activities with their friends. However, according to a recent research synthesis, many of these digital behaviors serve the same purpose and encompass the same core qualities as face-to-face ...

When Internet addiction is actually a good thing - The Washington Post

<https://www.washingtonpost.com/.../when-internet-addiction-is-actually-a-good-thing/> ▼

Dec 31, 2014 - When Internet addiction is actually a good thing. A new study from researchers at the University of Hong Kong claims that 420 million people are addicted to the Internet, about 6 percent of the world's population.

Technology can seriously damage your health - NaturalNews.com

https://www.naturalnews.com/033611_technology_health_effects.html ▼

Sep 19, 2011 - (NaturalNews) Did you know that technology can seriously damage your health? Tragically most people still don't get it. The World Health Organisation (W.H.O.) states very clearly that there are serious health risks from current exposure levels to electronic fields coming from wi-fi, laptops, mobile phones ...

How Smartphones Are Making Kids Unhappy - NPR

<https://www.npr.org/sections/health.../08/.../how-smartphones-are-making-kids-unhappy>

Aug 7, 2017 - Psychologist Jean Twenge has observed dramatic shifts in behavior among children who go through adolescence with smartphones. They're spending less time with friends and reporting greater anxiety.

Blame Society, Not the Screen Time - NYTimes.com

<https://www.nytimes.com/roomfordebate/...a.../blame-society-not-the-screen-time> ▼

Danah Boyd is a principal researcher at Microsoft Research, the founder and president of Data & Society and author of "It's Complicated: The Social Lives of Networked Teens." She is on Twitter. Updated July 11, 2016, 6:12 PM. Even though multiple generations have now grown up glued to the flickering light of the TV, we ...

Technology is changing the Millennial brain | PublicSource

<https://www.publicsource.org/technology-is-changing-the-millennial-brain/> ▼

Nov 14, 2015 - The U.S. Census Bureau says there are 83.1 million people between the ages of 18 to 34 in the nation. Studies show the average Millennial spends 18 hours per day using any time of digital media. And, 90 percent of young adults use social media, which is up from 12 percent in 2005, the Pew Research ...

The Many Social Benefits of Playing Video Games | LevelSkip

<https://levelskip.com> › Miscellaneous ▼

Dec 20, 2017 - Conventional wisdom suggests that children and adults who spend a lot of their free time playing video games are socially inept and that video games may stunt the social development of children. But is this true? Recent research suggests that video games may have many social benefits and that certain ...

FIGURE 3

Digital library designed to resemble a Google search page.

removing the comment section). The documents varied by type (e.g., blog post, newspaper article, popular magazine), source credibility, content trustworthiness, and perspectives on the topic

(see Table 1). Text length varied from 340 to 1,787 words. To verify the features of the documents in the digital library, the first and fourth authors independently coded each for the level

TABLE 1 Details of the documents in the digital library.

Title	Type	Source	Year	Length	Topic position*	Source credibility	Content credibility
An Ugly Toll of Technology: Impatience and Forgetfulness	Newspaper	<i>New York Times</i>	2010	859	Agree	High	High
Teenagers are Better Behaved and Less Hedonistic Nowadays	Newspaper	<i>The Economist</i>	2017	1787	Neutral	High	High
Secrecy is Dead. Here's What Happens Next.	Magazine	<i>Wired</i>	2017	1235	Agree	High	Low
Teens' Online Friendship Just as Meaningful as Face-to-Face Ones	Research press release	<i>Science Daily</i>	2017	340	Disagree	High	High
When Internet Addiction is Actually a Good Thing	Newspaper	<i>Washington Post</i>	2014	863	Disagree	High	Low
Technology Can Seriously Damage Your Health	Blog	<i>Natural News</i>	2011	631	Agree	Low	Low
How Smartphones Are Making Kids Unhappy	Website (Radio station)	<i>NPR</i>	2017	723	Agree	High	Low
Blame Society, Not the Screen Time	Newspaper	<i>New York Times</i>	2016	655	Disagree	High	Low
Technology is Changing the Millennial Brain	Blog	<i>Public Source</i>	2015	1438	Agree	Low	Low
The Many Social Benefits of Playing Video Games	Blog	<i>Levelskip</i>	2017	1269	Disagree	Low	High

*Topic position refers to the position presented in the document vis-à-vis the controversial statement: "Students today are overly dependent on technology to the detriment of their social, physical, emotional, and academic well-being."

of source credibility (high or low), overall content trustworthiness (high or low), and topic stance (agree, disagree, or neutral). The interrater agreement was 96.7%. Consensus was reached on all these dimensions of text features through discussion.

Sources judged as high in credibility were from reputable publishers or websites known for accurate and reliable reporting of information (e.g., the New York Times and The Economist). Sources low in credibility were from less well-established or personal outlets (e.g., PublicSource and Levelskip), and sites known for propagating pseudoscientific information (e.g., Natural News). Among the ten documents in the library, seven were judged as high in source credibility and three were rated as low in credibility.

As for content trustworthiness, the content of documents was considered trustworthy if the author presented relevant, accurate, and objective evidence to support their claim, and if the evidence was communicated in a logical and rigorous manner. Four documents were in the high content credibility category. In contrast, documents were considered to present low-credibility content if the authors made vague or unsupported arguments, presented claims without citations, or based the claims on personal experience and questionable evidence. Six documents were classified as low in content credibility.

Finally, with regard to the stances represented in the documents, four focused on the harmful effects of technology. For example, the article by NPR “How Smartphones are Making Kids Unhappy,” argued for the deleterious impact of smartphones on children’s socioemotional wellbeing. Five documents forwarded a positive view of technology. For example, the Washington Post article, “When Internet Addiction is Actually a Good Thing,” viewed high Internet addiction rates as a sign of socioeconomic improvements. One document, “Teenagers are Better Behaved and Less Hedonistic Nowadays” published by The Economist, showed a neutral stance by presenting evidence that supported both positive and negative sides of technology.

Procedure

The students completed the MSU task independently on their laptops. The task consisted of four components: (a) pre-reading questions, (b) digital library reading, (c) post-reading questions, and (d) argumentative essay writing. The first three components were completed on Qualtrics®, while the essays were composed in Microsoft® Word. First, students provided consent to participate in the study and completed the pre-reading

questions about their demographic information and their initial positions on the controversial topic. Specifically, students were presented with the topic statement and were asked to indicate the extent to which they agreed or disagreed with that statement on a 0 to 100 scale (0 = strongly disagree, 100 = strongly agree). They were then directed to record their position on the controversial topic and to provide a brief justification for that position.

Next, students proceeded to navigate the digital library and select documents for reading. Students were instructed to scan the search result page and read the linked documents. They were required to select at least four documents to read in depth. As they read the selected documents, they could highlight any information that stood out to them as particularly trustworthy or questionable and rate the overall credibility and usefulness of each document at the end of the page on scales of 0 to 100. The highlighting and document ratings, which were not part of this analysis, were elements of a separate project investigating students' source and content evaluations (Sun et al., 2020). Students were also allowed to take notes while they read.

After reading, students responded to the post-reading questions about their final positions on the topic. For these questions, they were again presented with the topic statement and were asked to indicate whether or not their position had changed and the extent to which they now agreed or disagreed with the statement on a 0–100 scale.

Finally, students composed their argumentative essays in a Word document. They were presented with the claim statement and were asked to follow the stated directions:

*Clearly **state** your position and write an argumentative essay regarding the viewpoint presented above. **Explain and justify** your position with sound reasoning.*

The students could refer to their notes as they wrote but could not re-access the digital library. They were instructed to write as much as they needed to articulate their arguments, but no specific length requirements were given.

Students were given instructions about each component of the MSU task in class and completed the task as an assignment outside of class. They were told to find a quiet place to first complete the Qualtrics[®] portion of the task (i.e., pre-reading questions, research in digital library, and post-reading questions) in one sitting and then write the essay as required. There were no time constraints on their reading, writing, or question responses, and all task components were completed within a 5-day period. For this study, we focused only on the argumentative essays for a Bayesian network analysis.

Argumentative essay scoring

The argumentative essays were scored based on a researcher-developed rubric that consisted of two sets of parameters: (a) adherence to task requirements, which included five key parameters for a multiple-source-based argumentative essay task (i.e., stating a *claim*, presenting *justifications*, referencing multiple *sources*, discussing *counterarguments*, and demonstrating adequate *writing ability*); and (b) integration of multiple sources, which consisted

of three core components for producing a well-integrated written essay (i.e., *critical analysis* of sources and contents, *synthesis* of multiple documents, and *overall cohesion* of information presentation). Each component was scored on a 0 to 2 scale, with a total possible score ranging from 0 to 16 (see Table 2 for rubric details).

Specifically, in terms of adherence to task parameters, effective written products were expected to meet the following criteria. First, the essay should present a clear *claim* reflecting the student's position on the controversial topic. A score of 2 was awarded when there was a clearly identifiable and well-articulated claim statement. A score of 1 was given when the claim was vaguely worded, while a 0 was given when a position statement is absent. Second, the essay included information from multiple *sources* from the digital library. A score of 2 was given when the essay cited at least two sources. A score of 1 was given if only one source was cited, whereas a score of 0 was assigned if no source was referenced.

Third, effective essays presented well-elaborated *justifications* for the claims. A score of 2 was awarded when students presented multiple supporting points for their overarching claims and fully discussed or substantiated those points with examples and evidence (e.g., research findings, data). When students did not fully elaborate or substantiate their supporting arguments, a score of 1 was awarded. Finally, A score of 0 was given when there were no supporting points or details provided or no clear connection between the claim and the supporting details.

Fourth, *counterarguments* representing contrasting or alternative views to students' claims were well articulated and fully addressed. For a score of 2, students needed to present more than one counterpoint to their claims and thoroughly discuss the counterevidence or counterevidence. When an essay only briefly mentioned a potential alternative or opposing view without elaboration, a score of 1 was given. A score of 0 was assigned if no counterpoints or counterevidence were addressed.

Finally, a well-crafted essay should manifest adequate *writing ability* that enables idea articulation. A score of 2 was awarded if the essay followed the mechanics of writing and, therefore, was fully comprehensible and coherent at the linguistic level. A score of 1 was given if the writing was comprehensible in general but was only moderately coherent. No credit was awarded if the essay was incomprehensible and incoherent.

As for components of multiple source integration, students' argumentative essays were assessed according to the following criteria. First, *critical analysis* was reflected in students' appraisal of source and content credibility or evaluation of the soundness of the arguments presented in the source documents. Strong critical analysis, warranting a score of 2, was evidenced when students critiqued authors' arguments based on the evidence provided (e.g., identifying that a causal relation cannot be inferred from the correlational data) or when they questioned the trustworthiness of the source to invalidate authors' arguments. In weaker cases of critical analysis that warranted a score of 1, students attempted at analyzing the information from the sources or evaluating authors' views or arguments, but such analyses or evaluations were superficial and unelaborated. Finally, a score of 0 was given if the essay did not demonstrate any evidence of analysis or critique of the sources or their contents.

TABLE 2 Rubric for scoring the argumentative essays on task and integration components.

Component	Score awarded		
	0	1	2
Claim	No claim statement, or an incomprehensible claim	Claim statement presented, but not fully articulated	Claim statement well-articulated
Sources	No source cited	Only one source cited	Multiple sources cited
Justification	No evidence of justifying how the evidence supports the claim	Some evidence of justification for the claim but not fully elaborated	Clear evidence of well-elaborated justification for the claim
Counterargument	No presence of counterargument(s)	Only one counterargument presented, or multiple counterarguments vaguely presented	Multiple counterarguments well-articulated
Writing ability	Incomprehensible paragraphs, or lexical incoherence	Generally comprehensible and moderately coherent lexically	Fully comprehensible, and highly coherent lexically
Critical analysis	No analysis of the sources or information from the sources	Limited analysis of information from sources, or superficial treatment of views or content in the sources	Strong critical analysis of the sources and the views, content, or information from the sources
Synthesis	No evidence of synthesis of information from across the sources	Some evidence of synthesis of information from across the sources in only part of the essay, or in a limited way	Clear evidence of synthesis of information from multiple sources at paragraph level or document level
Overall Cohesion	Disconnected or isolated ideas across paragraphs	Some evidence of connecting ideas within or between paragraphs	Strong connection within and between paragraphs; ideas flow naturally from one to another

The second key integration component, *synthesis*, was assessed based on the degree to which students meaningfully consolidated information from multiple sources in making their arguments. In strong cases of synthesis warranting a score of 2, students wove multiple pieces of information from different sources around their arguments, often within several paragraphs or across the entire document. Such synthesis could manifest when students pulled together research findings from two sources that supported the same point, or when they pointed out conflicts between information in two documents. In contrast, weaker evidence of synthesis (warranting a score of 1) was observed when a student included pieces of information from different sources in a loosely connected fashion or only in part of the essay. Further, when no connection between cited sources was identified, a score of 0 was given.

Lastly, *overall cohesion* was evidenced by the degrees of logical connection and flow of ideas within and between paragraphs. In highly cohesive essays (a score of 2), paragraphs were well organized and strongly connected, with clear transitions from one idea to the next, often indicated by connective words and phrases such as “however”, “therefore”, “further”, and “on the contrary.” A score of 1 was awarded if the essay demonstrated weak organizational structure of the ideas across the document or limited flow within or between paragraphs. A point of 0 was given when ideas were presented in a disconnected fashion.

Three independent raters evaluated a randomly selected 10.6% of the essays, yielding an intraclass correlation coefficient (ICC) of 0.93 for interrater reliability on the overall score. The ICCs for individual task parameters were: 0.76 for claim, 0.90 for counterargument, 0.85 for justification, 0.98 for sources, and 0.89 for writing ability. The ICCs for the integration components were: 0.85 for critical analysis, 1.0 for overall cohesion, and 0.85 for synthesis. Prior to scoring, the raters underwent a training process

that began with rubric familiarization. They were provided detailed rubrics for each criterion, which explained the scoring scales and what constituted each level of performance. The training then progressed to calibration sessions, where raters scored sample essays together in group settings. During these sessions, they discussed their rationales and worked to resolve any discrepancies, thereby aligning their understanding of the rubrics. After the training, the raters independently scored the essays.

Data analysis

The data from scoring the task and integration components were used to determine which of the two models more accurately reproduced the argumentative essay writing process with multiple documents using Bayesian network analysis. Each of the eight components that made up the nodes of the Bayesian networks was scored at three levels from 0–2 based on the rubric. Due to insufficient cases of participants who received the lowest score on *writing ability* ($n = 3$), stating a *claim* ($n = 1$), and presenting *justifications* ($n = 2$), the cases were combined with those who received a score of 1. Consequently, these components in the models had only two levels of performance.

Prior to fitting the data to the models, we considered the use of informative priors, which in Bayesian analysis are probability distributions that incorporate existing knowledge about the parameters before observing the data. While previous argumentative writing research exists, it primarily uses non-Bayesian methods making it challenging to translate directly into informative priors. Given this limitation, we opted for weakly informative priors, assigning uniform distribution of students’ probabilities of performing at different levels for each component. A uniform distribution, in this context, means that we assigned

equal initial probabilities to each performance level, rather than assuming that some levels were more likely than others. Subsequent studies can use information from this research to inform the selection of priors.

After fitting the models to the data, we evaluated how well the predictions made by the models matched the observed data using the leave-one-out cross-validation method. Next, the selected model was used to determine the most crucial components of writing an argumentative essay using multiple documents using the Bayesian Network belief updating procedure.

Transparency and openness statement

The data used to fit the Bayesian network model is available here https://osf.io/2jh3k/?view_only=d05f1a63bd794f63be960a84e8bd95ee. All other data associated with the study, methods used in the analysis, and materials used to conduct the research will be made available for research purposes upon reasonable request to the corresponding author.

GeNIe Modeller (BayesFusion, LLC) was used for all Bayesian Network modeling described in this study.

Results and discussion

Descriptive summary of student performance

Based on the scoring rubric for the argumentative essay, we determined that the mean performance for the eight components was 10.01 ($SD = 3.41$) as presented in Table 3. For components specific to the writing task, the students tended to score between 1 and 2 on the components: writing ability ($M = 1.44$, $SD = 0.55$), claim ($M = 1.88$, $SD = 0.36$), sources ($M = 1.55$, $SD = 0.76$), and justification ($M = 1.55$, $SD = 0.54$). However, a sizeable number of students were unable to provide a counterargument ($M = 0.90$, $SD = 0.84$).

This contrast may suggest that these undergraduates were unaware of the role of counterarguments in well-crafted argumentative essay. Alternatively, such a pattern may indicate that these students were operating under the belief that their goal was to “win” an argument and that excluding counterarguments would weaken their stance (Brown and Renshaw, 2000; Gilbert, 1997). Relatedly, this frequent absence of counterarguments could reflect myside bias or confirmation bias (Mercier, 2016; Stanovich et al., 2013) in which individuals tend to favor information that confirms their beliefs and disfavor information that counters them.

To gain a more comprehensive understanding of this phenomenon, we conducted a qualitative analysis of the notes from students who performed poorly on the counterargument component. Interestingly, we observed instances where students had engaged with multiple sources during the execution stage, producing high-quality notes, yet failed to incorporate this information into their essays during the production stage. Conversely, students who primarily focused on a single source demonstrated greater proficiency in producing counterarguments. While this qualitative analysis was not the primary focus of our

study, it suggests that the struggle with counter argumentation might be partially attributed to the challenges of managing multiple sources of information.

It is also possible that critical analysis of the sources and content is a prerequisite for producing a well-formed counterargument. We tested this possibility in the interconnected Bayesian network model by specifying a causal link from *critical analysis* to *counterargument*. The Bayesian Network model permitted us to test the direction and strength of this causal relation. We observed that students in our sample performed poorly on critical analysis ($M = 0.65$, $SD = 0.82$), which could explain the low performance on the counterargument component.

Students demonstrated superior performance on the synthesis component ($M = 1.12$, $SD = 0.76$) compared to both critical analysis and overall cohesion ($M = 0.90$, $SD = 0.67$) within the integration construct. This relative strength in synthesis may stem from students' experience with MSU assignments in college courses. Additionally, the task design, which instructed students to take notes on selected articles before composing argumentative essays, likely contributed to high synthesis scores. This goal-directed note-taking and review process has been shown to enhance encoding and learning (Kobayashi, 2006), potentially facilitating synthesis. However, the observed discrepancy between synthesis performance and counterargument generation highlights the intricate cognitive processes involved in crafting argumentative essays from multiple documents.

RQ 1: theoretical models of argumentative essay writing

The interconnected model (Model B in Figure 2) better captured the process of writing an argumentative essay from multiple documents compared to the linear model (Model A in Figure 2). It is worth noting that both models predicted students' performance levels above chance, but the results from the leave-one-out cross-validation procedure indicated that the interconnected model demonstrated higher overall prediction accuracy (65%) than the linear model (63%).

The primary difference between the two models was the interconnections among the various components. In the linear model, enactment of the task components (i.e., source, claim, justification, counterargument) led to synthesis, which gave way to critical analysis and overall cohesion. On the other hand, in the interconnected model, synthesis was a precondition for the successful use of sources, and critical analysis was needed for providing justifications and compelling counterarguments.

The results favoring the interconnected model indicate that a crucial part of the argumentative essay writing process unfolds in the execution stage as outlined by the IF-MT. In this stage, students select credible sources, extract important points and supporting details within and across documents, find associations, and prepare a mental or physical organization of what they have read. Therefore, synthesis and critical analysis appear to be crucial preconditions for enacting the task components in the essay writing process.

It must be noted that while both models demonstrated high prediction accuracy overall, the interconnected model exhibited

TABLE 3 Descriptive statistics for performance on argumentative essay.

Components	Levels of Performance (<i>n</i> = 105)						<i>M</i> (<i>SD</i>)
	0		1		2		
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
Writing Ability	3	2.86	53	50.48	49	46.67	1.44 (0.55)
Claim	1	0.95	11	10.48	93	88.57	1.88 (0.36)
Sources	17	16.19	13	12.38	75	71.43	1.55 (0.76)
Justification	2	1.90	42	40.00	61	58.10	1.56 (0.54)
Counterargument	42	40.00	31	29.52	32	30.48	0.90 (0.84)
Critical Analysis	59	56.19	23	21.90	23	21.90	0.65 (0.82)
Synthesis	24	22.86	44	41.90	37	35.24	1.12 (0.76)
Overall Cohesion	29	27.62	57	54.29	19	18.10	0.90 (0.67)
Total Score							10.01 (3.41)

several advantages in predicting individual task and integration components that we detail below.

Interconnected model outperformed linear model in predicting performance on task and integration components

The interconnected model correctly predicted the performance of students on task components 72.81% of the time, and its combined prediction accuracy rate for the integration components was 57.78%. In comparison, the linear model was 70.8% correct for task components and 55.23% for the integration components. Although the integration components' prediction accuracies were lower than the task components in both models, they were significantly higher than the prediction accuracy rates by chance (33.3%). [Figure 4](#) presents the prediction error rates for the two models for each of the task and the integration components. The prediction error rate is an inverse of prediction accuracy; so high accuracy and low error indicates a good-fitting model. It is vital to examine the prediction accuracy by task components and the three score levels ranging from 0 to 2 to unpack the granularity of differences between the two models.

Comparing the two models' prediction accuracy for specific task components showed that both performed equally well for the claim and justification components. The primary difference between the models occurred for the components of sources and counterarguments. The linear model could not predict performance on the source component associated with scores 0 and 1. On the other hand, the interconnected model, where synthesis was specified as a precondition for sources, successfully predicted the performance of those scoring 0 and 2 on the source component (lowest and highest scores possible) with an accuracy rate of 76.47 and 92%, respectively.

For counterargument, the linear model predicted performance levels 0 and 1 with accuracy rates of 83.34 and 53.12%, respectively, but could not predict the performance of those at score level 2. On the other hand, the interconnected model accurately predicted the highest performance level 2 (65.62%) and level 0 (76.19%), but struggled with the middle score of 1. The difference between the

two models was that the interconnected model specified critical analysis as a precondition for counterargument. It appears that when critical analysis is entered as a prerequisite for counter argumentation, the model performed better at predicting both the lack of counterarguments and exhibiting the highest competence at counter argumentation, but not the intermediate level. This indicates that critical analysis is not associated with only briefly mentioning a counterargument.

Overall, the interconnected model showed high accuracy for each component, but consistently failed to predict performance at score level 1 for claim, sources, justification, and counterargument, with accuracy rates of 0% for most and 50% for justification. Even critical analysis had a low accuracy rate of 13% for level 1. This difficulty in predicting intermediate performance aligns with our earlier observation about counterarguments. It suggests that level 1 performance, which often represents partial or developing skills, has a different relationship with other components than either high (level 2) or low (level 0) performance and the structure of a model fails to capture those relationships.

The interconnected model's struggle with level 1 predictions across components might suggest that the progression from basic to advanced skills in argumentative writing with multiple sources is not straightforward, making intermediate stages particularly challenging to model accurately alongside other performance levels. See [Supplementary material](#) for more information comparing the two models.

RQ 2: relative importance of task and integration components

To address the second research question, we used Bayesian network updating, a method that allows for estimating the probabilities of predefined hypothetical scenarios. Since the interconnected model demonstrated superior performance, we used this model to investigate the relative importance of various components. For instance, in one hypothetical scenario, we forecasted the likelihood of achieving different performance levels

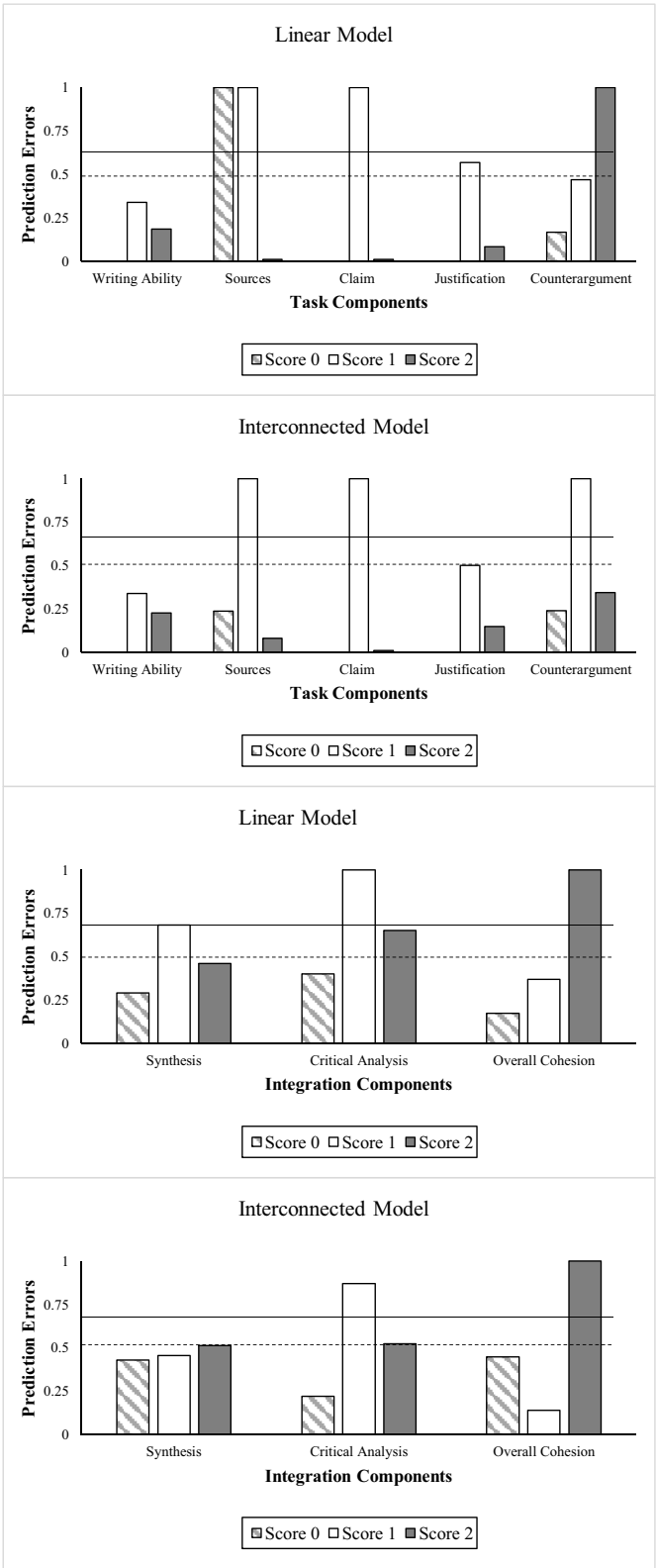


FIGURE 4
Plots showing prediction errors for task and integration components for linear and interconnected models. The dotted lines represent the prediction error rate by chance. The dotted line at 0.5 is for Writing ability, Claim, and Justification and the solid line at 0.67 is for all other components with three levels of measurement.

across all components when an individual excels in critical analysis. In the Bayesian network framework, this is analogous to probing which components contribute most to achieving the highest level of critical analysis and how this proficiency influences other components connected to it. We carried out this analysis for all task and integration components, synthesizing the results to gain insights.

In the interconnected Bayesian network, we began modeling with the assumption that students had an equal chance of performing at each of the three levels across different components before any observations were made, using what is known as an uninformative prior. This meant that for components with two levels, like writing ability, there was a 50-50 chance for each level, and for components with three levels, like sources, there was an equal chance of being at any of the three levels (33.33%). Next, the student performance data were used to estimate the probability distribution for each component. Finally, we estimated the key task and integration components based on the model using Bayesian network updating procedures.

Components were categorized as crucial to the integration process if achieving high-level performance (i.e., 100% probability of being at level 2) on that component increased the likelihood of performing well (i.e., being at level 2) on other components in the model. In contrast to other statistical tools, Bayesian network updating allowed us to observe the effect of change in performance on one variable on all other variables that serve as a cause or a consequence of that focal variable in modeling a process.

Writing ability and source use influenced the process, claim formulation showed no effect

In the Bayesian network model, the writing ability component headed the process and was connected to all the remaining task and integration components. Given the written nature of this task, this structure modeled that proficient writing should be a prerequisite skill for manifestation of task (source, claim, justification, and counterargument) and integration (synthesis, critical analysis, and overall cohesion) components that underlie the production of a written argumentative essay. However, our analysis revealed that writing competently (i.e., achieving a score of 2 with 100% probability) directly influenced only one of the task components, justification, and one integration component, synthesis. Specifically, we observed that excelling at the writing ability component increased the probability of also excelling in the corresponding highest levels for justification and synthesis to 75 and 52%, respectively (refer to Figure 5).

Source use strongly impacted the justification component, and source use was itself influenced by synthesis ability. Specifically, the likelihood of achieving a score of 2 on the justification component increased from 55 to 71% when performing with a 100% probability at level 2 on the sources component (see Figure 5). Notably, the integration component synthesis influenced the ability to incorporate multiple sources, a relationship supported by the student performance data. In a hypothetical scenario where synthesis performance was fixed at 100% for score level 2, a corresponding increase was observed in the ability to use multiple

sources, rising from a 68% chance to achieve score level 2 to 91% (refer to Figure 5).

The Bayesian updating procedure showed that making a claim was independent of writing ability, and a well-formulated claim was not a prerequisite for sources, justification, or counterargument (see Figure 6). The data from the essays indicated that most undergraduates were able to formulate a well-articulated claim. The controversial topic in this study was selected by a group of undergraduates demographically similar to those who participated in this study. As a result, students appeared to have had an opinion on the topic regardless of the multiple viewpoints presented in the documents. When asked to indicate if their opinion changed due to reading the multiple documents presented in the library, 84% of the students responded, “No, my position has not changed.” This result, therefore, could be an artifact of the fact that most students had an opinion on this topic before they read the topic and simply articulated that opinion in their claim statement. The relations specified by the model should be tested with a topic where students do not have strong preconceived notions.

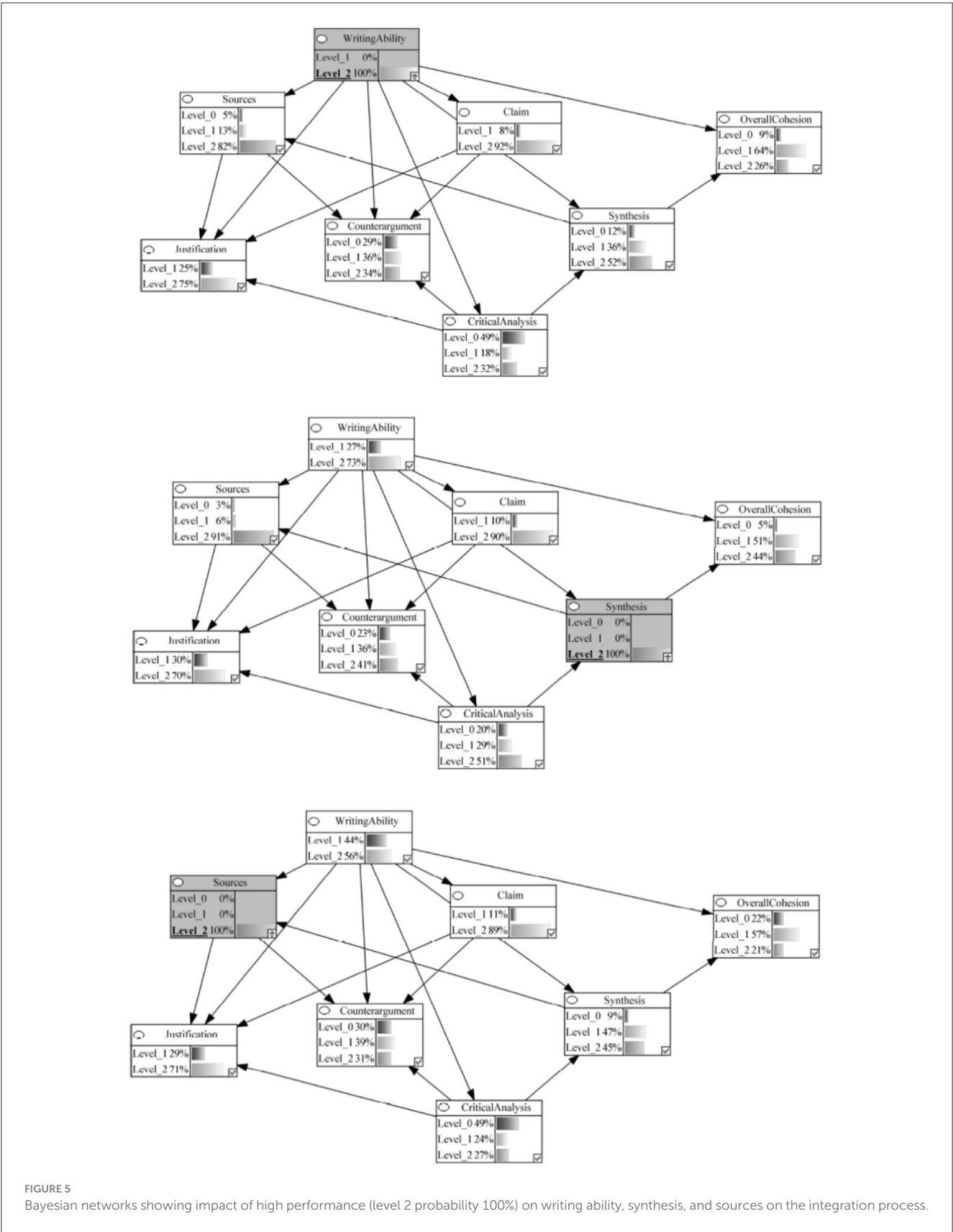
Critical analysis had a substantial impact on task and integration components

The model specified that critical analysis undergirds the ability to present justifications, construct counterarguments, and synthesize information from multiple documents. This pivotal nature of critical analysis was evidenced through Bayesian updating, with the probability of justification increasing to 74%, of counterargument from 27 to 46%, and synthesis from 33 to 78% for being at score level 2 when critical analysis performance was at 100% for the highest score level (see Figure 6).

According to the IF-MT, successful implementation of justification and counterargument components occurs not only at the production stage but begins early on in the process when students analyze the requirements of the argumentative essay task in the preparation stage (see Figure 1). In the argumentative essay, the manifestation of critical analysis appears to be a continuous trace of a critical analytic stance adopted early on during the preparation stage and enacted during execution and production (Sun et al., 2020).

Synthesis ability was pivotal to the argumentative writing process

We observed that when students achieved highest points on synthesis (score level 2), the probability of referring to multiple sources in their argumentative essays increased dramatically from 68 to 91%. The Bayesian updating in our network demonstrates that knowing a student has strong synthesis skills allows us to make much more confident predictions about their use of multiple sources in argumentative writing. It is important to note that this pattern also reveals that some students (9%) achieved high synthesis while not referring to multiple sources, and many students (68%) referenced multiple sources without achieving high synthesis scores, underscoring the complexity of the probabilistic relationship between synthesis and multiple source use. Students



might successfully synthesize information using fewer sources in some cases, and conversely, students might reference multiple sources without effectively synthesizing the information.

Despite this complexity, our findings align with the Documents Model Framework (Britt and Rouet, 2012; Perfetti et al., 1999) and the IF-MT. According to Documents Model Framework, readers

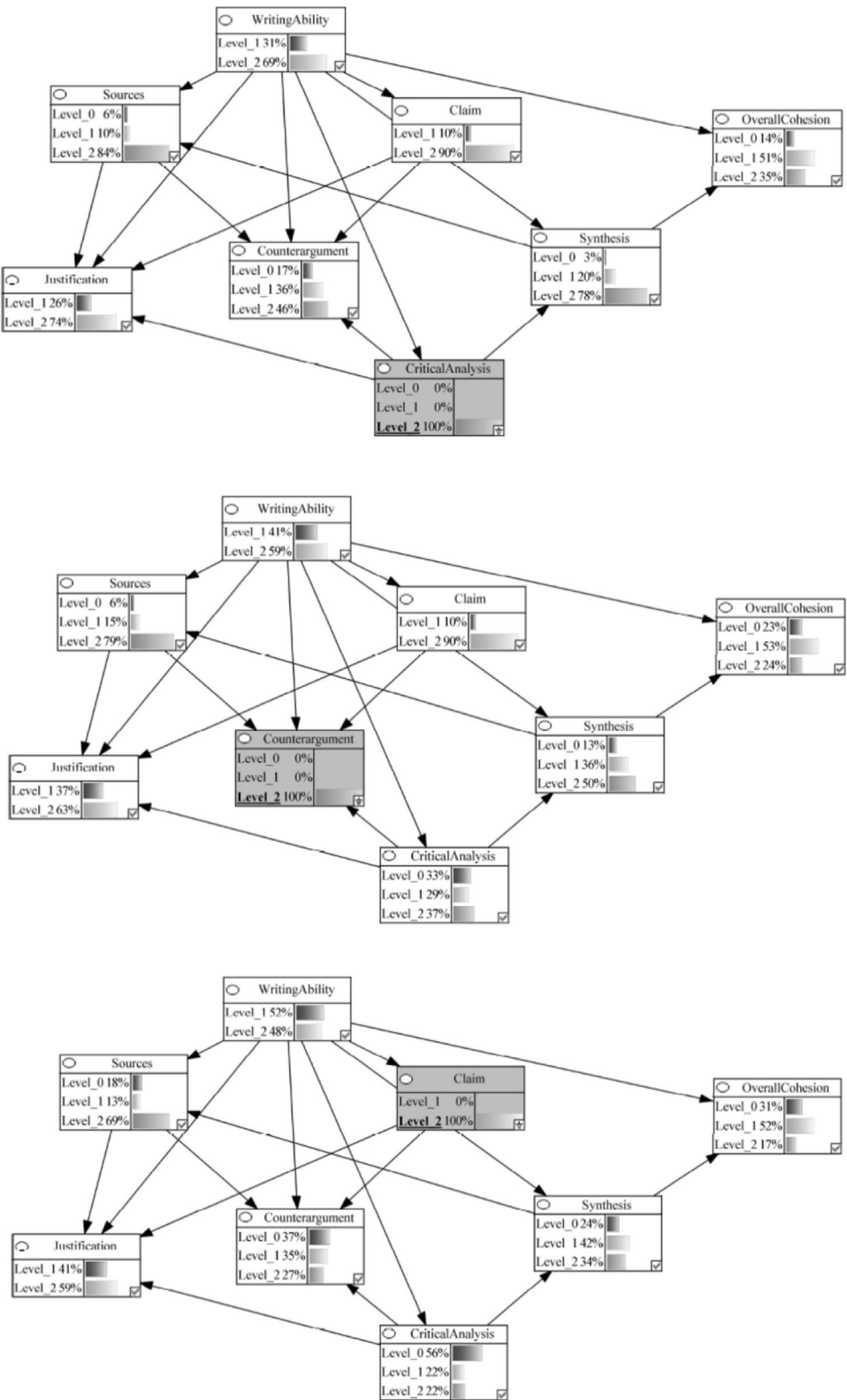


FIGURE 6
Bayesian networks showing impact of high performance (Level 2 Probability 100%) on critical analysis, counterargument, and claim on the integration process.

constructing mental representations from multiple documents develop both a situation model (representing document content) and an intertext model (representing source information and connections among documents). Both our empirical results and the theoretical framework establish synthesis as a predictor of students' ability to effectively draw from and reference multiple sources in their writing.

The influence of synthesis extends beyond source use to other aspects of writing quality. For instance, successful synthesis more than doubled the likelihood of producing highly cohesive essays, with overall cohesion rates rising from 17 to 44%. This demonstrates that the cognitive skill of building connections across texts directly translates into measurable improvements in written composition.

Similarly, synthesis showed strong connections to argumentative quality. Previous investigations focusing on synthesis have also attested to the importance of inter-textual relations in forming evidence-backed opinions on controversial topics (e.g., Kobayashi, 2009). In the present study, we found that students can produce strong counterarguments (i.e., achieving a score of 2) when they have at least 50% probability of achieving the highest performance level in synthesis. This is in contrast to the baseline probability of 33% for achieving a score of 2 in synthesis, which represents the chance probability if scores were randomly distributed across the three possible levels (0, 1, and 2) with equal likelihood (refer to Figure 6).

Synthesis, which is the opposite of piecemealing, and a manifestation of the ability to draw intra- and inter-textual connections in documents emerged as an indispensable competency to support integration in the MSU argumentative essay task.

Conclusion and implications

The aims of this study were multifaceted. We set out to model the process of written argumentation using Bayesian network analysis in the context of a multiple source use task. We tested the comparative prediction accuracy of two theoretically viable models—the linear model and the interconnected model. The models were constructed based on argumentation and multiple source use literatures. Although both models made better than chance predictions, the interconnected model reproduced the data with higher accuracy than the linear model. After selecting the higher-performing model, we used Bayesian updating as an innovative method to pinpoint the key components in the argumentative writing process in MSU contexts. The insights gleaned from this analysis can be used to inform instruction of argumentative writing in the internet age, where students have to contend with an informational deluge.

Implications for research: modeling the componential process of argumentative essay writing

This study is among the first attempts to model the process of writing an argumentative essay using Bayesian network analysis.

Bayesian network analysis is distinct from other modeling tools because it provides information on student performance in probabilistic terms and allows for testing causal connections. The evidence supporting the interconnected model indicated that the integration components, especially critical analysis, are a key driver of the argumentative writing process. This finding lends credence to the stage-based framework (IF-MT) proposed by List and Alexander (2019), wherein essential cognitive actions are undertaken well before the production of the essay. They forward that *preparation* and *execution* are crucial stages before *production*. It is in these earlier two stages that we see the enactment of task analysis, building of intra- and inter-textual links, and critical analysis of the sources and the content contained in them. But the manifestation of these earlier actions is readily available in the *production* stage, where students actually produce the written essay. The strength of the modeling procedure used in this study is that we were able to gain insights into causal links among the components that undergird the process of writing by assessing the product.

The Bayesian network modeling approach we presented in this study can be flexibly adapted to model other MSU writing tasks. The models we tested were rooted in the production phase of the IF-MT and incorporated both task-specific and general integration components. The hybrid Bayesian network modeling approach, blending theory with tasks with computational learning, provides a framework for researchers and practitioners to investigate other writing task processes. One possible avenue for future MSU investigations would be to retain the integration components while adjusting the task parameters to suit other types of writing tasks.

Bayesian analysis becomes even more powerful as we gather additional information and build on previous investigations. In this study, we did not have any prior information on the performance of students. However, now we have data about performance on each component, for example, we know that most undergraduates can provide a claim statement but struggle with counterarguments. Future research studies can use more informative priors by drawing from the current research to improve the predictive power of Bayesian networks. In the [Supplementary material](#), we provide complete conditional probability tables for each node in our network, which researchers can directly incorporate as Dirichlet priors when studying similar populations. For educational practice and intervention research, we can design instruction that supports the needs of specific types of students with increasing precision by collecting more data.

Implications for practice: supporting integration in an MSU argumentative essay task

Our study on undergraduate argumentative essay writing using multiple documents revealed several key components that students struggle with. Critical analysis emerged as a crucial element, with its causal connection to counterargument playing a pivotal role in improving the overall integration process. We focused on these aspects because they underlie effective decision-making, problem-solving, and functioning in democratic societies, as noted by scholars like Dewey (1933) and Diamond (2013).

TABLE 4 Forms of relational reasoning.

Form	Definition	Example in argumentation context
Analogy	Recognizing meaningful similarities	Identifying when two sources present parallel arguments
Anomaly	Identifying deviations from patterns	Recognizing when evidence contradicts an established pattern
Antinomy	Recognizing mutual exclusivity	Understanding when accepting one position necessitates rejecting another
Antithesis	Identifying direct oppositions	Recognizing when sources directly contradict each other

Additionally, the ability to synthesize information surfaced as another critical parameter.

The findings indicated that while undergraduates can provide justifications given adequate writing skills, their major weakness lies in considering and rebutting counter views. This shortcoming is significant because constructing a reasoned argument requires more than mere justifications; it demands the consideration of multiple perspectives and the evaluation of evidence supporting contrasting views. To address this, we propose that interventions or instruction supporting counter argumentation should focus on honing students’ critical analysis skills. One avenue for doing this would be through training in relational reasoning.

Relational reasoning, the ability to discern patterns in information streams, encompasses four distinct forms: analogical, anomalous, antinomous, and antithetical (see Table 4). By developing these skills, students could draw deeper connections among multiple documents, identifying similarities, dissimilarities, and contradictions across texts. This enhanced ability would prepare them to critically analyze complex information, synthesize ideas from various sources, and develop counterarguments, ultimately improving their overall argumentation skills.

Our study methodology, which required students to compose essays using only their notes without direct access to digital texts, potentially encouraged deeper engagement with source materials during the initial reading phase. This approach may have led to more thorough note-taking and enhanced synthesis scores. However, the significant discrepancy observed between synthesis and counter argumentation performance suggests that even with this potentially beneficial note-taking process, students still struggled with more complex argumentative task components.

In real-life settings, where students would likely have access to original texts, we might expect some differences in performance. While direct text referencing might improve accuracy, it could potentially reduce the depth of engagement that our task structure encouraged. Educators applying our findings should consider incorporating strategies that combine the benefits of note-taking with the practical reality of text availability. This could include teaching effective annotation and quick-reference techniques, ensuring that students develop both deep engagement with texts and

practical skills for managing multiple sources in their argumentative writing.

Limitations and future directions

A primary limitation of the current study is the use of a single task topic and one essay per student, which potentially restricts the external validity of our findings. Additionally, to reflect the three stages of the IF-MT framework, we directed students to first make notes and then use only those notes to craft their essays without referring back to the original sources or revising their work. While this controlled approach aligned with our theoretical framework and may have enhanced synthesis skills, it deviated from authentic writing practices where students typically consult source materials throughout the writing process and revise their essays. Future studies should consider incorporating a more diverse range of topics, allow students to consult articles during writing, and provide opportunities for revision to better represent how argumentative writing naturally unfolds in academic contexts.

While our study design prioritized ecological validity by allowing students to complete the assignment as homework, this approach introduced a tradeoff with experimental control and may have contributed to variability that a controlled laboratory setting could have minimized. Despite this limitation, the homework format better reflects the conditions under which students typically complete writing assignments. The compromises between naturalistic conditions and procedural control that we made highlight the complex challenges in studying writing processes in classroom settings.

Turning to the central aspect of our study, a key limitation of our modeling approach was that we only compared two alternative Bayesian network models. While our methodology and task instructions (e.g., requiring students to forward a claim and use only their notes to write essays) necessarily imposed certain constraints on component sequences, additional plausible models exist that we did not test. For instance, models that modify individual connections between components in our Models A and B, or models incorporating direct causal influences on claim formulation, might offer alternative explanations of the data. The *apriori* models we tested were theoretically driven, based on the literature on argumentation and multiple source use. However, we acknowledge that argumentative writing involves complex processes that could be represented through various network structures. For instance, we recognize the complexity in relations between specific components, such as synthesis and multiple source use, where alternative causal directions may be plausible—it could be that exposure to multiple sources enables better synthesis, rather than synthesis driving multiple source use, or that students might successfully synthesize with fewer sources while others might reference multiple sources without effectively synthesizing the information. Future research should systematically test a broader range of alternative models, including those with different directional relationships between components, to strengthen causal inferences about argumentative writing processes.

Another limitation of our approach is that it does not fully capture the entire writing process as laid out by the IF-MT. Specifically, the Preparation and Execution stages were not

incorporated into the Bayesian network analysis. Our working assumption was that we would see traces of the processes that occur during the previous stages in the written essay (Production stage). However, we found as part of unplanned analysis that there was a disconnect in the quality of the notes and the quality of the essays. Future research should aim to include these crucial stages to provide a more comprehensive understanding of the writing process.

The IF-MT, drawing from the cognitive affective engagement model (CAEM; List and Alexander, 2017), articulates that students can adopt four stances toward task completion—disengaged, affectively engaged, evaluative, critical analytic. According to the IF-MT, all four of these stances combine affective and cognitive aspects that are not just preexisting individual attributes but also iterate within the context of the task. It follows that if educators are interested in developing critical analytic orientation in students, instruction should consider the role of affective dimensions such as engagement in and motivation for the task in addition to developing cognitive resources. The current study does not measure affective involvement. However, we concede that this dimension also plays a role in successfully integrating multiple documents in a goal-directed context. Further, it is essential to note that epistemic beliefs are also significant predictors of the processing of multiple documents (Bråten et al., 2011; Ferguson et al., 2013), but we did not include them in our Bayesian model. Future investigations should study the distinct contribution of affect and epistemic beliefs in integration and how they undergird the higher-order cognitive skills, such as, critical analysis.

Final word

The 21st-century context is riddled with open-ended, ill-structured problems in information-rich digital spaces. One of the pressing challenges that educational researchers must respond to is how to foster the habits of mind that equip students to integrate relevant information from multiple sources to solve a problem. In this investigation, we have demonstrated an innovative, adaptive, and theoretically driven method for modeling a written task using multiple documents of variable source and content credibility. Using this method, we not only shed light on how the componential process unfolds but also determined vital areas where students require support. Subsequent researchers using Bayesian analysis should build on the findings of this study to further enhance our understanding of the process of writing in MSU contexts and the challenges that students face.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Institutional Review Board University of Maryland. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. YS: Conceptualization, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. PA: Conceptualization, Methodology, Writing – review & editing. HZ: Investigation, Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1560088/full#supplementary-material>

References

- Alexander, P. A., and the Disciplined Reading and Learning Research Laboratory (2020). "Relational reasoning: the bedrock of integration within and across multiple representations, documents, and perspectives," in *Handbook of Learning from Multiple Representations and Perspectives*, eds. P. van Meter, A. List, D. Lombardi, and P. Kendeou (New York, NY: Routledge), 401–424. doi: 10.4324/9780429443961-26
- Alexander, P. A., and the Disciplined Reading and Learning Research Laboratory. (2012). Reading into the future: competence for the 21st century. *Educ. Psychol.* 47, 259–280. doi: 10.1080/00461520.2012.722511
- Almond, R. G., Mislevy, R. J., Steinberg, L. S., Yan, D., and Williamson, D. M. (2015). *Bayesian Networks in Educational Assessment*. New York, NY: Springer. doi: 10.1007/978-1-4939-2125-6
- Anmarkrud, Ø., McCrudden, M. T., Bråten, I., and Strømso, H. I. (2013). Task-oriented reading of multiple documents: Online comprehension processes and offline products. *Instr. Sci.* 41, 873–894. doi: 10.1007/s11251-013-9263-8
- Aristotle, 4th century BCE (2019). *The Art of Rhetoric* (R. Bartlett, Trans.). Chicago, IL: The University of Chicago Press.
- Asterhan, C. S. C., and Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *J. Educ. Psychol.* 99, 626–639. doi: 10.1037/0022-0663.99.3.626
- Asterhan, C. S. C., and Schwarz, B. B. (2016). Argumentation for learning: well-trodden paths and unexplored territories. *Educ. Psychol.* 51, 164–187. doi: 10.1080/00461520.2016.1155458
- Barzilai, S., Tal-Savir, D., Abed, F., Mor-Hagani, S., and Zohar, A. R. (2021). Mapping multiple documents: from constructing multiple document models to argumentative writing. *Read. Writ.* 36, 1–39. doi: 10.1007/s11145-021-10208-8
- Bereiter, C., and Scardamalia, M. (1987). *The Psychology of Written Composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bråten, I., Britt, M. A., Strømso, H. I., and Rouet, J. F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: toward an integrated model. *Educ. Psychol.* 46, 48–70. doi: 10.1080/00461520.2011.538647
- Braasch, J. L., and Bråten, I. (2017). The discrepancy-induced source comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educ. Psychol.* 52, 167–181. doi: 10.1080/00461520.2017.1323219
- Braasch, J. L. G., Bråten, I., Britt, M. A., Steffens, B., and Strømso, H. I. (2014). "Sensitivity to inaccurate argumentation in health news articles: potential contributions of readers' topic and epistemic beliefs," in *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences*, eds. D. N. Rapp J. L. G. Braasch (The MIT Press), 117–137. doi: 10.7551/mitpress/9737.003.0009
- Brand-Gruwel, S., Wopereis, I., and Walraven, A. (2009). A descriptive model of information problem solving while using Internet. *Comp. Educ.* 53, 1207–1217. doi: 10.1016/j.compedu.2009.06.004
- Britt, M. A., and Rouet, J. F. (2012). Learning with multiple documents: component skills and their acquisition," in *Enhancing the Quality of Learning: Dispositions, Instruction, and Learning Processes*, eds. J. R. Kirby and M. J. Lawson (Cambridge: Cambridge University Press), 276–314. doi: 10.1017/CBO9781139048224.017
- Brown, R. A. J., and Renshaw, O. D. (2000). "Collective argumentation: a sociocultural approach to reframing classroom teaching and learning," in *Social Interaction in Learning and Instruction: The Meaning of Discourse for the Construction of Knowledge*, eds. H. Cowie and G. van der Aalsvoort (Amsterdam: Elsevier Science), 52–66.
- Chambliss, M. J., and Murphy, P. K. (2002). Fourth and fifth graders representing the argument structure in written texts. *Discourse Processes* 34, 91–115. doi: 10.1207/S15326950DP3401_4
- De La Paz, S. (2005). Effects of historical reasoning instruction and writing strategy mastery in culturally and academically diverse middle school classrooms. *J. Educ. Psychol.* 97, 139–156. doi: 10.1037/0022-0663.97.2.139
- De La Paz, S., and Felton, M. K. (2010). Reading and writing from multiple source documents in history: effects of strategy instruction with low to average high school writers. *Contemp. Educ. Psychol.* 35, 174–192. doi: 10.1016/j.cedpsych.2010.03.001
- Dewey, J. (1933). *How We Think*. Boston, MA: Dover Publications.
- Diamond, A. (2013). Executive functions. *Ann. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750
- Ecker, U. K., and Antonio, L. M. (2021). Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect. *Mem. Cognit.* 49, 631–644. doi: 10.3758/s13421-020-01129-y
- Ferguson, L. E., Bråten, I., Strømso, H. I., and Anmarkrud, Ø. (2013). Epistemic beliefs and comprehension in the context of reading multiple documents: examining the role of conflict. *Int. J. Educ. Res.* 62, 100–114. doi: 10.1016/j.ijer.2013.07.001
- Ferretti, R. P., and Fan, Y. (2016). "Argumentative writing," in *Handbook of Writing Research, 2nd Edn*, eds. C. A. MacArthur, S. Graham, J. Fitzgerald (New York, NY: Guilford Press), 301–315.
- Flower, L., and Hayes, J. R. (1980). The cognition of discovery: defining a rhetorical problem. *Coll. Compos. Commun.* 31, 21–32. doi: 10.58680/coc198015963
- Freedman, A., and Pringle, I. (1984). Why students can't write arguments. *English Educ.* 18, 73–84. doi: 10.1111/j.1754-8845.1984.tb00668.x
- Galbraith, D. (1998). "Writing as knowledge-constituting process," in *Studies in Writing: Vol. 4 Knowing What to Write: Conceptual Processes in Text Production*, eds. G. Rijlaarsdam, E. Esperet, M. Torrance, and D. Galbraith, (Amsterdam: Amsterdam University Press), 139–159.
- Galbraith, D., and Torrance, M. (1998). "Conceptual processes in writing: From problem solving to text production," in *Studies in Writing: Vol. 4 Knowing What to Write: Conceptual Processes in Text Production*, eds. G. Rijlaarsdam, E. Esperet, M. Torrance, and D. Galbraith (Amsterdam: Amsterdam University Press), 1–12.
- Gilbert, M. (1997). *Coalescent Argument*. Mahwah, NJ: Erlbaum.
- Graham, S. (2018). A revised writer (s)-within-community model of writing. *Educ. Psychol.* 53, 258–279. doi: 10.1080/00461520.2018.1481406
- Gutmann, A. (1999). *Democratic Education (Revised edition)*. Princeton, NJ: Princeton University Press. doi: 10.1515/9781400822911
- Hayes, J. R., and Flower, L. S. (1980). "Identifying the organization of writing processes," in *Cognitive Processes in Writing*, eds. L. W. Gregg and E. R. Steinberg (Hillsdale, NJ: Lawrence Erlbaum Associates), 3–30.
- Hess, D., and Avery, P. (2008). "Discussion of controversial issues as a form and goal of democratic education," in *The SAGE Handbook of Education for Citizenship and Democracy*, eds. J. Arthur, I. Davies, and C. Hahn (London: SAGE), 506–518. doi: 10.4135/9781849200486.n40
- Jensen, F. V. (1996). *An Introduction to Bayesian Networks*. London: UCL Press.
- Jensen, F. V., and Nielsen, T. D. (2007). *Bayesian Networks and Decision Graphs*. New York, NY: Springer. doi: 10.1007/978-0-387-68282-2
- Jiménez-Aleixandre, M. P., and Erduran, S. (2007). "Argumentation in science education: an overview," in *Argumentation in Science Education: Perspectives from Classroom-Based Research*, eds. S. Erduran, and M. P. Jiménez-Aleixandre (Dordrecht: Springer Science), 3–27.
- Kobayashi, K. (2006). Combined effects of note-taking/-reviewing on learning and the enhancement through interventions: a meta-analytic review. *Educ. Psychol.* 26, 459–477. doi: 10.1080/01443410500342070
- Kobayashi, K. (2009). Comprehension of relations among controversial texts: effects of external strategy use. *Instr. Sci.* 37, 311–324. doi: 10.1007/s11251-007-9041-6
- Kuhn, D. (1991). *The Skills of Argument*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511571350
- Kuhn, D., and Modrek, A. S. (2021). Choose your evidence. *Sci. Educ.* 31, 21–31. doi: 10.1007/s11191-021-00209-y
- Lauritzen, S. L., and Spiegelhalter, D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *J. R. Stat. Soc. Ser. B.* 50, 157–224. doi: 10.1111/j.2517-6161.1988.tb01721.x
- Leitão, S. (2003). Evaluating and selecting counterarguments: studies of children's rhetorical awareness. *Written Commun.* 20, 269–306. doi: 10.1177/0741088303257507
- List, A., and Alexander, P. A. (2017). Cognitive affective engagement model of multiple source use. *Educ. Psychol.* 52, 182–199. doi: 10.1080/00461520.2017.1329014
- List, A., and Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *Educ. Psychol.* 54, 20–39. doi: 10.1080/00461520.2018.1505514
- List, A., Du, H., and Lyu, B. (2022). Examining undergraduates' text-based evidence identification, evaluation, and use. *Read. Writ.* 35, 1059–1089. doi: 10.1007/s11145-021-10219-5
- Luna, M., Villalón, R., Martínez-Álvarez, I., and Mateos, M. (2022). Online interventions to help college students to improve the degree of integration of their argumentative synthesis. *Read. Writ.* 36, 937–963. doi: 10.1007/s11145-021-10248-0
- Mason, L., and Scirica, F. (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learn. Instr.* 16, 492–509. doi: 10.1016/j.learninstruc.2006.09.007
- Mateos, M., Martín, E., Cuevas, I., Villalón, R., Martínez, I., and González-Lamas, J. (2018). Improving written argumentative synthesis by teaching the integration of conflicting information from multiple sources. *Cogn. Instruct.* 36, 119–138. doi: 10.1080/07370008.2018.1425300
- McCann, T. M. (1989). Student argumentative writing knowledge and ability at three grade levels. *Res. Teach. English*, 62–76. doi: 10.58680/rte198915528

- McClelland, J. L., Rumelhart, D. E., and Hinton, G. E. (1986). "The appeal of parallel distributed processing," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, eds. D. E. Rumelhart, J. L. McClelland, and the PDP Research Group (Cambridge, MA: MIT Press), 3–40. doi: 10.7551/mitpress/5236.001.0001
- McGrew, S. (2021). Skipping the source and checking the contents: an in-depth look at students' approaches to web evaluation. *Comp. Sch.* 38, 75–97. doi: 10.1080/07380569.2021.1912541
- Means, M. L., and Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognit. Instr.* 14, 139–178. doi: 10.1207/s1532690xc1402_1
- Mercier, H. (2016). "Confirmation bias-myside bias," in *Cognitive Illusions* (New York, NY: Psychology Press), 109–124. doi: 10.4324/9781315696935-11
- Murphy, K. (1998). *A Brief Introduction to Graphical Models and Bayesian Networks*. San Francisco, CA: Morgan Kaufmann.
- National Assessment of Educational Progress (1994, 1999, 2002). *NAEP Writing Report Card*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.
- National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standards (English Language Arts)*. Washington D.C.: National Governors Association Center for Best Practices and Council of Chief State School Officers.
- Nussbaum, E. M. (2011). Argumentation, dialogue theory, and probability modeling: alternative frameworks for argumentation research in education. *Educ. Psychol.* 46, 84–106. doi: 10.1080/00461520.2011.558816
- Nussbaum, E. M., and Schraw, G. (2007). Promoting argument-counterargument integration in student writing. *J. Exp. Educ.* 76, 59–92. doi: 10.3200/JEXE.76.1.59-92
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann. doi: 10.1016/B978-0-08-051489-5.50008-4
- Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.
- Pearl, J., and Russell, S. (2003). "Bayesian networks," in *Handbook of Brain Theory and Neural Networks*, ed. M. A. Arbib (Cambridge, MA: MIT Press), 157–160. Available online at: https://ftp.cs.ucla.edu/pub/stat_ser/r216-reprint.pdf
- Pearson, K. (1904). *Mathematical Contributions to the Theory of Evolution*. London: Dulau and Co.
- Perelman, C., and Olbrechts-Tyteca, L. (1969). *The New Rhetoric: A Treatise on Argumentation* (Transl. by J. Wilkinson, and P. Weaver). Notre Dame, IN: University of Notre Dame Press.
- Perfetti, C. A., Rouet, J.-F., and Britt, M. A. (1999). "Toward a theory of documents representation," in *The Construction of Mental Representations During Reading*, eds. H. van Oostendorp and S. R. Goldman (Mahwah, NJ: Lawrence Erlbaum Associates), 99–122.
- Rouet, J. F. (2006). *The Skills of Document Use: From Text Comprehension to Web-Based Learning*. Mahwah, NJ: Erlbaum.
- Sinharay, S. (2006). Model diagnostics for Bayesian networks. *J. Educ. Behav. Stat.* 31, 1–34. doi: 10.3102/10769986031001001
- Sparks, J. R., and Rapp, D. N. (2011). Reader's reliance on source credibility in the service of comprehension. *J. Exp. Psychol. Learn. Mem. Cognit.* 37, 230–247. doi: 10.1037/a0021331
- Stadtler, M. (2017). The art of reading in a knowledge society: commentary on the special issue on models of multiple text comprehension. *Educ. Psychol.* 52, 225–231. doi: 10.1080/00461520.2017.1322969
- Stanovich, K. E., West, R. F., and Toplak, M. E. (2013). Myside bias, rational thinking, and intelligence. *Curr. Directions Psychol. Sci.* 22, 259–264. doi: 10.1177/0963721413480174
- Sun, S., Alexander, P. A., Singh, A., and Zhao, H. (2020). *Characterizing College Students' Multiple Source Use in an Argumentative Essay Task: A Profile Analysis [Poster]*. Washington, DC: APA 2020 Convention.
- Ting, K. M. (2017). "Confusion matrix," in *Encyclopedia of Machine Learning and Data Mining*, eds. C. Sammut and G. I. Webb (Boston, MA: Springer). doi: 10.1007/978-1-4899-7687-1_50
- van Boekel, M., Lassonde, K. A., O'Brien, E. J., and Kendeou, P. (2017). Source credibility and the processing of refutation texts. *Mem. Cognit.* 45, 168–181. doi: 10.3758/s13421-016-0649-0
- van Eemeren, F. H. (2013). In what sense do modern argumentation theories relate to Aristotle? The case of pragma-dialectics. *Argumentation* 27, 49–70. doi: 10.1007/s10503-012-9277-4
- van Eemeren, F. H., and Grootendorst, R. (2004). *A Systematic Theory of Argumentation: The Pragma-Dialectical Approach*. Cambridge: Cambridge University Press.
- van Eemeren, F. H., Grootendorst, R., and Henkemans, F. S. (1996). *Fundamentals of Argumentation Theory: A Handbook of Historical Backgrounds and Contemporary Developments*. Mahwah, NJ: Lawrence Erlbaum Associates. doi: 10.2307/358423
- Van Wijk, D. (1998). "Conceptual processes in argumentation: a developmental perspective," in *Studies in Writing: Vol. 4 Knowing What to Write: Conceptual Processes in Text Production*, eds. G. Rijlaarsdam, E. Esperet, M. Torrance, and D. Galbraith (Amsterdam University Press), 31–47.
- Vandermeulen, N., van den Broek, B., Van Steendam, E., and Rijlaarsdam, G. (2020). In search of an effective source use pattern for writing argumentative and informative synthesis texts. *Read. Writ.* 33, 239–266. doi: 10.1007/s11145-019-09958-3
- Walton, D. (2007). *Media Argumentation: Dialectic, Persuasion and Rhetoric*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511619311
- Wiley, J., and Voss, J. F. (1999). Constructing arguments from multiple sources: tasks that promote understanding and not just memory for text. *J. Educ. Psychol.* 91, 301–311. doi: 10.1037/0022-0663.91.2.301
- Yan, D., Mislevy, R. J., and Almond, R. G. (2003). Design and analysis in a cognitive assessment (ETS RR-03-32). *ETS Res. Rep. Ser.* 2003, i–47. doi: 10.1002/j.2333-8504.2003.tb01924.x



OPEN ACCESS

EDITED BY

Rui Alexandre Alves,
University of Porto, Portugal

REVIEWED BY

Ana Camacho,
University of Porto, Portugal
Golnar Mazdayasna,
Yazd University, Iran

Xiao Tan,
Utah State University, United States

*CORRESPONDENCE

Talal Musaed Alghizzi
✉ tmalghizzi@imamu.edu.sa

RECEIVED 20 August 2024

ACCEPTED 22 July 2025

PUBLISHED 13 August 2025

CITATION

Abdel Latif MMM, Alghizzi TM and
Alshahrani TM (2025) Teacher motivational
strategies in Saudi university EFL writing
classes: a qualitative study.
Front. Psychol. 16:1483456.
doi: 10.3389/fpsyg.2025.1483456

COPYRIGHT

© 2025 Abdel Latif, Alghizzi and Alshahrani.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Teacher motivational strategies in Saudi university EFL writing classes: a qualitative study

Muhammad M. M. Abdel Latif¹, Talal Musaed Alghizzi^{2*} and
Tahani Munahi Alshahrani²

¹Faculty of Graduate Studies of Education, Cairo University, Giza, Egypt, ²College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

Empirically, motivating students to write is an issue yet to be given due research attention. Some previous relevant works have suggested guidelines for motivating writing students, but the studies qualitatively exploring the realities of writing teachers' use of motivational strategies remain scant. In this study, we investigated Saudi university teachers' perceptions of their students' English writing de-motivation symptoms (i.e., signs or indicators) and causes of lack of motivation to write, and the ways they motivate students to write and participate in classroom activities. We explored these issues through using interviews with 33 teachers (17 males and 16 females) who had English writing instruction experiences at five Saudi universities. The 33 teachers identified seven main symptoms of students' writing de-motivation (procrastinating assignment submission, engaging rarely in classroom activities, showing writing apprehension, copying others' writing, skipping classes, perceiving writing value negatively, and experiencing writing block), and they referred to five causes of it (students' poor language and writing ability, uninteresting topics, ineffective teaching, previous poor experiences, and the cognitive nature of writing). The teachers also reported using eight main motivational strategies in their English writing classes. For these teachers, class size is a very influential factor in their use of motivational strategies. The results generally suggest that writing motivation is yet to be given more attention in Saudi university English writing classes. The study provides the following recommendations: fostering teacher motivation literacy, activating the use of motivational strategies in writing classes, and minimizing class size.

KEYWORDS

writing motivation, writing de-motivation, motivational strategies, writing teacher, L2 writing, Saudi universities

1 Introduction

Second language (L2) students' motivation plays an important role in their language acquisition and learning. Therefore, due attention should be paid to nurturing L2 students' motivation. The task of motivating students to write is a much more complicated one than motivating them to learn a language (Abdel Latif, 2021). In classes covering multiple L2 areas, it is normal to find more than one teacher instructing students, and thus assuming the responsibility of motivating them to learn the target language. In writing classes, only the writing teacher is responsible for motivating students to write. What makes motivating students to write a more challenging task for the teacher is the fact that writing is the most cognitive of all the language skills as it requires much more time and deeper mental processes.

Despite its importance, the issue of how teachers motivate their students to write has been rarely researched. There have been a few relevant empirical studies on the realities of using

motivation strategies (Cheung, 2018; Lee and Lin, 2022; Mali, 2017; Rosina, 2017; Saranraj et al., 2014). Meanwhile, other published relevant works have only provided guidelines for motivational strategies in writing classes (e.g., Abdel Latif, 2019a, 2019b, 2021; Bruning and Horn, 2000; Reeves, 1997; Troia et al., 2012; Walker, 2003). Thus, such obvious scarcity requires conducting further research on teachers' use of motivational strategies in L2 writing classes. Specifically, we need to understand writing teachers' perceptions of their students' de-motivation symptoms and sources, how they try to motivate them to write, and the factors influencing teachers' use of motivational strategies in writing classes. Understanding these issues could help in identifying what writing teachers need to make their instruction more motivating and how to help writing students avoid de-motivation symptoms. The present study attempted to tackle this under-explored research area by examining Saudi university English-as-a-foreign-language (EFL) teachers' perceptions of their students' writing de-motivation and its causes, the motivational strategies they use in writing classes, and the role of students' writing competence and classroom size as potential factors influencing their use of motivational strategies.

2 Writing motivation and teacher motivational strategies

Writing motivation is a multifaceted construct encompassing a number of sub-constructs. It can be generally defined as "learners' liking or disliking of writing situations and perceived value of writing, the situational feelings they experience while writing and the way they regulate them, the beliefs about their writing ability and skills, and their desired goals for learning to write" (Abdel Latif, 2021, p. 3). In light of this taxonomy, writing motivation constructs can be classified into the following four categories: (a) writers' attitudinal/dispositional feelings (writing apprehension, attitudes to writing, and the perceived value of writing); (b) situational perceptions and operations (writing anxiety and motivational regulation of writing, respectively); (c) self-ability beliefs (writing self-efficacy and self-concept); and (d) writing learning goals (i.e., mastery or task goals versus performance ones). See Abdel Latif (2019a, 2021) for detailed discussions of conceptualization and measurement issues and the framework of writing motivation constructs.

Literature indicates that L2 students' writing motivation is shaped by a number of factors. Collectively, these factors include: students' personal variables (i.e., gender, age and cultural background), their writing and language performance and beliefs, and learning and instruction practices and the issues related to them such as teaching materials and practices, and teacher and peer feedback (for more details, see Abdel Latif, 2021; Karaca and Inan, 2020; Pajares et al., 2007). Instructional practices particularly play an important role in motivating students' to write. According to Drew and Sørheim (2009), students' motivation greatly depends on instructional practices and teacher-student relationship.

Of particular relevance to the impact of instructional practices on language learners' motivation is the teachers' use of motivational strategies. Teachers' motivational strategies can be defined as the procedures used to generate, stimulate and maintain students' learning motivation (Dörnyei, 2001; Guilloteaux and Dörnyei, 2008). Dörnyei (2001) proposed a framework encompassing the following

components of motivational practices in L2 teaching: creating basic motivational conditions, generating students' initial motivation, maintaining and protecting their motivation, and encouraging their positive retrospective self-evaluation.

Regarding the frameworks or guidelines proposed for motivating students to write, a few works have dealt with this issue. Compared to L2 learning motivational strategies, the guidelines in these frameworks are of more specific nature as they relate to writing motivation only. Besides, some of these frameworks pertain to particular writing motivation constructs rather than others. Reeves (1997), for instance, suggested the following guidelines for minimizing students' writing apprehension: listening to fearful writers and conferring with them, varying writing modes, and preparing them for peer feedback activities. Walker (2003) also called for cultivating students' self-efficacy through involving them in choosing writing topics, encouraging their writing self-regulation and strategy use, providing them with self-evaluation opportunities, and employing learning-oriented assessment. Meanwhile, Troia et al. (2012) viewed that students' writing self-ability beliefs can be enhanced through fostering writing skill learnability and improveability beliefs whereas their writing learning mastery goals can be incentivized through prioritizing and modifying them and emphasizing effort attributions. Likewise, Limpo and Alves (2017) believed that teachers can improve students' writing self-ability beliefs and mastery goals via "proposing challenging and meaningful assignments, providing frequent opportunities for success, emphasizing the process of learning, stressing self-improvement over social comparisons, giving regular progress feedback, praising for effort rather than for ability, and promoting students' sense of autonomy" (pp. 118–119).

On the other hand, two more detailed frameworks for fostering students' writing motivation were provided by Bruning and Horn (2000) and Abdel Latif (2021). Bruning and Horn (2000) proposed a set of writing motivation procedures related to the following four guidelines: cultivating students' functional beliefs about writing, engaging them in performing authentic tasks, developing a supportive learning environment, and creating a motivating learning atmosphere. More recently, Abdel Latif (2019b, 2021) suggested six main guidelines for motivating L2 students to write; each guideline has a list of pedagogical procedures, totalling 42 ones for all the six guidelines. The six guidelines are: (a) nurturing and fostering students' writing motivational perceptions, beliefs and goals (7 pedagogical procedures); (b) using appropriate teaching materials and writing tasks (6 procedures); (c) meeting students' language and writing performance needs (6 procedures); (d) integrating technological tools in writing instruction (6 procedures); (e) optimizing teacher feedback (9 procedures), and (f) orchestrating peer assessment activities (8 procedures). Each guidelines with its pedagogical procedures can foster particular dimensions in students' writing motivation. These guidelines and procedures are not all used at one time but employing each depends on the stages of the writing course and students' needs. Figure 1 shows the six motivational guidelines proposed by Abdel Latif (2021).

As noted above, some guidelines have been proposed for motivating language learners and also for motivating students to write. Chronologically, early frameworks of the two types have almost occurred in the same period (for example, Dörnyei, 2001 versus Reeves, 1997; Bruning and Horn, 2000). Contrarily, early published frameworks of language teacher motivational strategies have

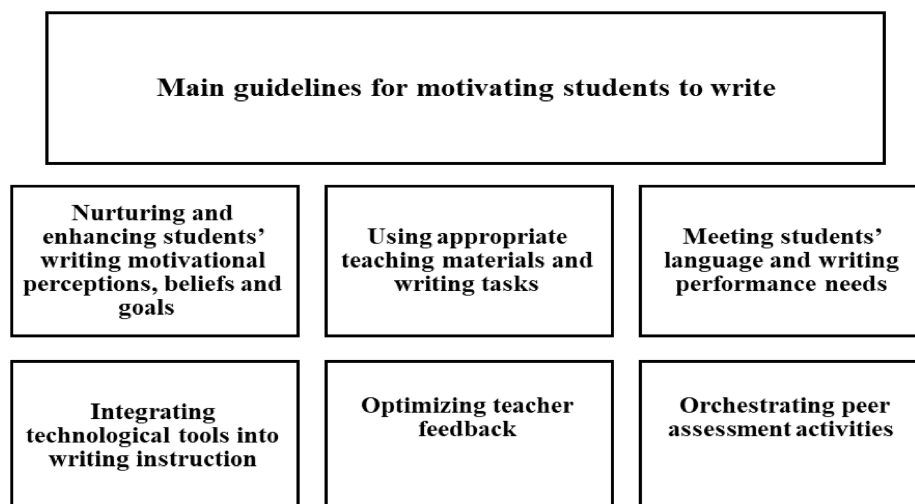


FIGURE 1

Abdel Latif's (2021) guidelines for motivating students to write (For the list of the 42 pedagogical procedures in the six guidelines, see Abdel Latif, 2021, pp. 150–151).

synchronizing been accompanied by empirical relevant studies (e.g., Dörnyei and Csizér, 1998) while research on writing teacher motivational strategies seems to have only occurred in the last decade. Since this writing motivational strategy research strand is rather recent, not many relevant studies have been published. This issue is further explained in the next section.

3 Previous studies

Previous studies on motivational strategies in writing classes are of three main categories: (a) interventional studies; (b) studies investigating students' perceptions of the motivational impact of writing instruction; and (c) studies dealing with the realities of writing teachers' use of motivational strategies. Many interventional studies of writing motivation have been published. The studies reported by Cruz Cordero et al. (2023), Zarrinabadi et al. (2023), Kim and Kim (2024), and Shen and Bai (2024) are examples of the recently published ones in this research strand. In their review of various issues in the early twenty-first century writing motivation research conducted in school settings, Camacho et al. (2021) highlighted the positive impact of some teaching treatments utilizing strategy instruction, collaborative writing, task type, and digital tools (e.g., blogs, wikis, games and web-based applications) on motivating students to write. Likewise, Abdel Latif (2021) reviewed writing studies experimenting the following six types of instructional treatments for motivating students to write: technology-supported writing instruction, writing strategy instruction, feedback provision techniques, genre-based instruction, writing task interest-based instruction, and therapeutic training. Of these instructional intervention types, technology-supported, strategy, genre, and task-interest-based instruction were specifically effective in developing students' writing motivation.

There have also been many studies on students' perceptions of the motivational impact of particular writing instruction types. Some of these studies revealed that students' writing demotivation may be developed as result of inappropriate instruction practices (Atay and

Kurt, 2006; Tsao et al., 2017), and the lack interesting teaching materials (Lo and Hyland, 2007). Other studies have dealt with the motivational impact of writing teacher feedback. One relevant large-scale study was reported by Yu et al. (2020) who explored how 1,190 Chinese university students perceived the impact of the following feedback types on their writing motivation: (a) scoring feedback given according to some descriptors; (b) process-oriented feedback given on multiple text drafts; (c) expressive feedback encompassing praise, criticisms and suggestions; (d) peer feedback; (e) students' self-evaluation of their own texts; and (f) written corrective feedback. Their participant students' writing motivation was found to be hindered by both process-oriented and written corrective feedback, but fostered by scoring, peer and self-feedback, and expressive feedback. These results emphasize the important role classroom feedback practices play in writing motivation.

Some other studies have explored writing students' perceptions of potential teacher motivational strategies. In a qualitative study, Mali (2017) used an open-ended questionnaire and sample lesson plans to explore 65 Indonesia university students' perceptions of teachers' instructional practices deemed motivating for them. Mali identified 120 questionnaire statements indicating the strategies the students perceived to be motivational in their writing classes. These strategies were related to: individualized material explanation and feedback, utilizing supportive technologies, engaging students in collaborative writing tasks and peer feedback, making jokes, playing songs, creating a friendly atmosphere with students, sharing learning strategies to students, and enabling students to be autonomous in their language learning. In another learner-centred study at a Hong Kong university, Lee and Lin (2022) investigated the motivational strategies teachers use in postgraduate English academic writing courses. They collected guided reflective pieces from 59 doctoral students who were asked to reflect upon the motivational strategies their teachers used regularly, and to describe the motivational impact of such strategies on them. This study showed that the writing teachers' effective motivational strategies as reported by their students include: using games and group work, using additional learning materials, giving students more

practice, raising students' awareness of their errors, and giving individualized instructional attention.

Only a few studies have been conducted on the realities of writing teachers' use of motivational strategies. These studies have been concerned with different international language learning environments. Saranraj et al. (2014), for instance, examined 19 teachers' motivational strategy use at an Indian university through using semi-structured interviews and a questionnaire with items assessing 27 instructional strategies. Their results showed that the motivational strategies the teachers reported using most frequently were raising students' awareness of the value and importance of the target activity, modeling task performance, and describing the task properly.

Two other studies on the realities of writing teacher motivation have combined teacher and student data. In a Norwegian high school, Rosina (2017) investigated teachers' motivational strategies in L2 writing classes through collecting questionnaire data from 100 students and conducting interview with three of their instructors. The teachers in Rosina's study had similar motivational strategies in their writing classes, and their strategies were associated with students' writing motivational levels. These teachers' motivational strategies include: using videos and visual aids, supporting students' writing through reading sources, providing students with positive and individualized feedback, and encouraging them through coursework marks. The teachers were also found unaware of more or less effective motivational strategies, and had difficulties in providing students with the needed motivational support due to time constraints. In the Singaporean higher education context, Cheung (2018) investigated the relationship between writing teachers' motivational strategies and students' motivation. The teachers taking part in this study were provided with a one-hour training in using motivational strategies in writing classes. The training was based on Dörnyei (2001)'s motivational strategy framework. The data was collected through surveying 344 students' perceptions, and observing 13 teachers in writing classes and surveying their practices. The observational data in Cheung's study showed that the strategies the teachers used pertained to generating task-specific motivation and maintaining it, and encouraging positive and retrospective self-evaluation. The results also revealed that the higher frequency of the teachers' reported use of strategies for generating students' initial classroom motivation was positively associated with students' positive attitude and high self-confidence.

Some points are noteworthy about the above-reviewed scarce studies on the realities of teacher use of writing strategies. First, the recent dates of publishing or reporting these few empirical works indicate that investigating writing teacher motivational strategies is a recent research strand. In contrast to this recent and scant research, teacher motivational strategies gained much earlier attention in general language learning motivation studies (e.g., Dörnyei, 2000, 2001; Dörnyei and Csizér, 1998). Overall, the above-reviewed studies can be regarded as initial attempts in exploring L2 writing teacher use of motivational strategies. Second, these studies have investigated motivational strategies in writing classes from different research angles. Third, interviews and questionnaires are two commonly used data sources in these studies, though qualitative data is relatively more common. Moreover, in developing data sources and analyzing data, the above studies have depended on general language learning motivational strategies frameworks (e.g., Dörnyei, 2001; Guilloteaux

and Dörnyei, 2008). Arguably, drawing upon the more relevant frameworks proposed for motivating students to write (e.g., Abdel Latif, 2021; Bruning and Horn, 2000) may reveal richer insights into writing teacher use and awareness of motivational strategies; an issue that seems to have been tackled only in language motivation research (e.g., Beshir, 2017; Waddington, 2018). Finally, previous studies also indicate that teacher use of motivational strategies is context-related. In other words, the use of such motivational strategies may vary from one context to another. As for the factors influencing writing teachers' use of motivational strategies, these are yet to be explored.

Since research on writing teacher motivational strategies is still in its infancy, further studies are needed in this strand to address the above-mentioned methodological and contextual gaps. The previous few studies are not without their limitations which have been mainly caused by the general motivational frameworks used and the nature of data collected. As a result, they have revealed a limited range of motivational strategies in writing instruction. Besides, the context-specific nature of writing teacher motivational strategies calls for exploring them in different international L2 settings. Accordingly, in-depth studies in this area could have important implications for improving L2 writing instruction and promoting students' writing motivation in specific language education environments. Their findings could also be of utmost importance to those interested in nurturing students' writing motivation, raising writing teachers' awareness of motivation strategies, and developing more robust survey instruments for assessing these strategies.

Taking the above-mentioned issues into account, the present study explored Saudi university EFL teachers' perceptions of their students' writing de-motivation and its causes and their use of motivational strategies in English writing classes. The study also investigated the potential impact of students' writing competence and class size on teachers' use of motivational strategies. Previous research suggests that writing competence correlates positively with students' writing motivation (e.g., Abdel Latif, 2015, 2021; Ferris, 2002; Karlen and Compagnoni, 2017; Teng et al., 2020). Likewise, general language education literature also implies that large class size could negatively influence teachers' ability to use motivational strategies (Cheng and Dörnyei, 2007; Dörnyei, 1998; Guilloteaux, 2013). This study seems to be the first attempt of its kind in addressing writing teacher motivational strategies in Saudi Arabia; recent writing motivation studies in this context have tackled issues other than instructional motivational strategies (e.g., Abdel Latif et al., 2024; Abdel Latif et al., 2024; Abdel Latif et al., 2024; Alsahil et al., 2024). Earlier L2 writing research in the Saudi context was mainly concerned with issues such as students' linguistic errors, rhetorical problems and writing processes, and their responses to particular instructional techniques (for a comprehensive review, see Abdel Latif, 2011). As it seems, the realities of writing teacher motivational strategies have hardly been given any research attention in Saudi Arabia. The originality and significance of the present study stems from its context which has unique cultural and educational characteristics. Issues such as single-gender education, and the nature of English writing instruction and difficulties at Saudi universities could differently shape students' writing de-/motivation and their teachers' motivational strategies. Accordingly, the unique contribution of the present study lies in offering insights into writing teacher motivational strategies from the Saudi context drawing upon a more detailed and relevant framework (Abdel Latif, 2019b, 2021).

4 The present study: research questions and method

As implied above, the present study tried to answer the following three research questions:

1. How do Saudi university EFL teachers perceive their students' writing de-motivation symptoms and sources?
2. How do these teachers try to motivate their students to write and participate in classroom writing activities?
3. To what extent do students' writing competence and class size influence teachers' use of motivational strategies?

We drew upon qualitative data to answer these research questions through using semi-structured interviews which enabled us to study the target issues from a more in-depth angle.

4.1 Research setting and participants

The study was conducted with a sample of faculty members who have taught writing to English majors at five Saudi universities. In the 4-year English language programmes these students attend at the five universities, English writing is taught as a core curriculum course over 4–5 terms depending on the study plan adopted by each college. In the multiple English writing courses taught, students learn writing different essay genres, including narrative, descriptive, opinion, and argumentative essays. The class size in writing courses relatively varies from one to another university, but according to the interviewees from the five universities it normally ranges from 20 to 35 students and it increases in female campuses in which a larger number of students study English as an academic major compared to male campuses.

Thirty-three faculty members took part in this study, 17 males and 16 females. They were teaching English writing at five Saudi universities (8 at University A, 7 at University B, 7 at University C, 6 at University D, and 5 at University E). In this study, we used the purposive sampling approach because we attempted to collect data from participant teachers with writing instruction experiences. We also decided to collect interview data from participants working at five universities as this would make the sample representative enough of writing teachers in the Saudi higher education system. An experience of teaching more than two writing courses was a pre-requisite for inviting the participants to take part in the study. Prior to starting the data collection process, the authors communicated with colleagues at the five universities to get a list of the faculty members teaching writing courses, and only those with the target writing instruction experience pre-requisite were invited through emails or phone calls to take part in the study. All the participants were PhD holders and they were of different academic ranks. The 33 teachers had varied teaching experiences ranging from three years to twenty-two years. They had also taught a number of English writing courses, ranging from 3 courses to more than 20 courses. With regard to their nationalities, the majority of the participants were Saudis ($n = 25$), and the other participants were: Egyptian (3), Jordanian (2), Sudanese (2), and Yemeni (1). All the participant teachers took part in the present study voluntarily and based on informed consent.

4.2 Semi-structured interviews

The present study made use of semi-structured interviews as its only data source because they allow a two-way communication mode between the researcher and interviewees, and thus help in understanding the what and why of the phenomenon investigated and allow raising follow-up questions about pertinent issues. We developed a set of guiding semi-structured interview questions in light of the research questions and the relevant literature. An expert language researcher read the guiding interview questions for face validity check, and confirmed they were appropriate for the research purpose and questions. These guiding questions focused on the teachers' language and writing instruction experiences, their conceptualizations of students' writing motivation, their perceptions of students' writing de-motivation symptoms (i.e., signs) and sources, the strategies they use for motivating students to write and participate in classroom activities, the potential influence of students' writing competence and class size upon their use of motivational strategies, and how the teachers associate students' writing motivation in their classes with the instructional procedures, teaching materials and topics used, technology use, and teacher and peer feedback (see the guiding interview questions in [Appendix 1](#)). In developing the guiding interview questions about the teachers' perceptions of students' writing de-motivation symptoms and sources and their common motivational strategies (questions 2–4), we tended to raise broad questions to gain insights into the teachers' actual writing de-motivation diagnosis practices and general de-motivation alleviating strategies. For the interview questions 5–11, we depended on reviewing literature on writing de-/motivation correlates and instructional motivational strategies.

4.3 Data collection and analysis

The data collection stage lasted for 6 weeks in which we obtained interview protocols from the 33 teachers who responded positively to our participation invitation. The semi-structured interviews were conducted individually with each participant. Both the second and third authors interviewed the male and female participants, respectively. The larger number of interviews ($n = 18$) were conducted on a face-to-face basis, whereas 10 participants were interviewed through online voice applications due to distance, and five other participants preferred to answer the interview questions in a written form in English. Since we were not able to raise follow-up questions in response to the written answers provided by five participants, this small interview portion has limitations in this regard; however, it added detailed information which helped us profile the teachers' writing de-motivation diagnosis and motivational strategies. In the face-to-face and online interviews, each participant teacher was interviewed in English or Arabic, or using a mixture of both languages, depending on their language preference. The interviews were guided by the questions we developed, and follow-up questions were also raised for eliciting the interviewees' pertinent opinions and narratives. All interviews lasted for 55–70 min.

The data analysis started with transcribing the interviews conducted in English, and translating the Arabic interviews and interview parts and then transcribing them in English. The translations of the Arabic interview parts were made by the second

and third authors, and reviewed and edited by the first author for meaning preservation verification; all the three authors are Arabic-native speakers with previous Arabic-English-Arabic translation experiences. Thus, we had all the 33 interviews transcribed in English. In our interview data analysis, we depended on the following procedures: exploring the data and initially categorizing it, identifying the descriptions related to each category in the interview protocols, reviewing and refining and initially identified categories, and confirming the evidence emerging from the data (Lodico et al., 2006). We read the interview protocols independently and categorized the emerging themes in them deductively using our research questions as broad guidelines and inductively through examining the sub-themes in each main category. Following this individual analysis, we met online to discuss the emerging themes each one of us identified. Through our online group discussion of the emerging themes in the individual analyses, we further revised and reconfigured them (Merriam, 1998), resolved the discrepancies in the data analysis, and reached agreed-upon labels for all the sub-themes. The trustworthiness of our interview data analysis was verified by an expert researcher who read four analyzed interview protocols to determine how much he would agree with the analysis made. The collaborator researcher had a very high agreement with the themes and categories we identified (93%) in the four protocols, and his comments were taken into account for refining some few dimensions in the data analysis. Guided by Abdel Latif (2021)'s writing motivation framework and motivational strategy guidelines, we organized these into categories related to the research questions.

5 Results of the study

In the following sub-sections, we present the results of the data analysis in light of the research questions. Each sub-section includes the answer of one research question.

5.1 The teachers' perceptions of students' writing de-motivation symptoms and sources

The teachers' interview answers showed they had varied conceptualizations of writing de-motivation symptoms or signs. Collectively, the 33 teachers identified the following seven symptoms of their students' writing de-motivation: procrastinating or skipping essay assignment submission, having a little engagement in classroom writing activities, showing a negative attitude toward writing, copying online materials or others' writing, skipping writing classes, having a low-perceived value of writing, and experiencing writing block. Table 1 provides the frequencies of these writing de-motivation symptoms as reported by the teachers, along with sample interview excerpts indicating them.

The frequencies of the symptoms imply how common the teachers have found them in their writing classes. As may be concluded, some teachers mentioned one symptom of students' writing de-motivation, while others referred to two or more. For

example, the teacher in the following interview excerpt is talking about several signs such as experiencing a negative attitude toward writing, skipping writing classes, and having little engagement in classroom activities:

Based on my teaching experience, I often identify students with a lack of writing motivation through certain behaviours. ... Occasionally they do not attend course classes, and they have a lack of interest in learning writing. ... They show a limited participation in topic discussions or group activities. ... And they also have frequent distractions or off-task behaviours during classroom activities. (Teacher 14)

Regardless of the number of de-motivation signs mentioned by each teacher, all the symptoms they gave mainly pertain to the attitudinal dimension of writing de-motivation, which includes students' negative attitudes toward writing, avoidance behaviors, and the perceived value of writing. In their description of the de-motivation sources, the teachers also talked about students' low language and writing ability beliefs, but they did not refer to any other signs or symptoms related to students' situational experiences (e.g., writing anxiety and low self-regulation), or the lack of writing achievement goals. Such limited conceptualization seems to have also negatively influenced the variety of motivational strategies the teachers use in their writing classes; this issue is discussed in the following subsection.

On the other hand, the interviews revealed important issues about the teachers' perceptions of their students' writing de-motivation sources. Overall, the 33 teachers identified five causes; these are: students' poor linguistic knowledge and writing ability, uninteresting writing topics, ineffective teaching materials and techniques, previous poor writing learning experiences, and the cognitive nature of writing tasks. Table 2 gives the frequencies of the teachers' mentions of these causes or sources and sample interview parts. Such frequencies suggest how often they occur in Saudi university writing classes.

On their descriptions of the role played by linguistic knowledge and writing ability levels, the interviewees ($n = 19$) indicated it is strongly associated with their students' writing motivation; as one interviewee explained:

Students with a good writing performance are always motivated to write. ... But students with a poor performance are usually less motivated. It is rare to find a high-level student low-motivated. If this happens, it is usually because the writing topic is less challenging and mediocre. (Teacher 23)

The majority of these 19 interviewees linked low-motivated students' poor writing with their inability to use appropriate vocabulary and grammar in their writing. Some other interviewees associated it with their inability to generate ideas and organize them even in their L1 (i.e., Arabic). For both teams of teachers, students' low-perceived level of English language proficiency and writing ability causes them not to participate actively in classroom writing activities because they fear to be criticized for their poor texts.

As for the influence of writing topics on students' motivation, the eight teachers highlighting this issue generally believed that a writing topic can be motivating to students if it is interesting, familiar to them,

TABLE 1 The teachers' conceptualizations of writing de-motivation symptoms.

The writing de-motivation symptom	No. of interviewees referring to it	Sample interview excerpt
Procrastinating or skipping essay assignment submissions	12	<i>I usually notice students' inadequate writing motivation when they ask me to extend the deadline or when they do not submit required essays. (Teacher 31)</i>
Having a little engagement in classroom activities	11	<i>When these students participate in writing activities or discussions, they pretend to do brainstorming when I walk past that group but in reality you can tell they are zoned out. (Teacher 20)</i>
Showing a negative attitude toward writing and its assignments.	8	<i>I typically notice low writing motivation in my writing classes when students consistently exhibit a lack of enthusiasm towards writing assignments... These students sometimes openly say they do not like writing or show signs of reluctance to start writing assignments. (Teacher 24)</i>
Copying online materials or others' writing	6	<i>They are the students who rely on looking at their classmates' writing and copying it. Sometimes they may also copy online essays or ask others to write essays for them. ... I can easily notice this in the essays they submit as there is a wide difference between the students' low levels and the high quality of the essays they submit. (Teacher 1)</i>
Skipping writing classes	5	<i>The students who are not motivated skip writing classes and do not attend them regularly. (Teacher 7)</i>
Having a low-perceived value of writing	4	<i>Some students perceive little relevance of writing skills in real life... This leads them to take writing courses only because they are mandatory, and not because they expect to benefit from them. (Teacher 19)</i>
Experiencing writing block	1	<i>They feel that writing is difficult and they do not know how to start the task, and so they have no aptitude to write in English. (Teacher 5)</i>

challenging enough and matches their levels. One interviewee also mentioned that topic interest can be a gender-related issue:

Female and male students may also react to some topics differently; for example, if female students are asked to write about sports or social issues, they will find this discouraging. (Teacher 18)

The seven teachers referring to the use of ineffective teaching materials and techniques as a potential cause of writing de-motivation thought that for writing learning materials and instruction to be motivating, they should not be boring, very difficult and should match students' interests. Meanwhile, these teachers had varied views regarding the perceived motivational impact of the writing teaching materials they use in their classes. Some teachers believed these teaching materials were motivating enough to students while others though they did not meet students' needs. Finally, a fewer number of teachers attributed students' writing de-motivation to learning previous experiences and the cognitive nature of writing tasks ($n = 4$ and 2 , respectively). While the first cause suggests students' long-term negative writing learning experiences have had a de-motivational impact on them, the second cause implies the cognitively demanding nature of the text composing process does not match some students' learning styles.

5.2 The teachers' instructional motivational strategies

The interviewed teachers reported using eight main strategies for motivating their students to write and to participate in classroom

activities. Table 3 shows these strategies, the number of the interviewed teachers referring to them, and related sample interview excerpts. The motivational strategies the teachers reported using are: optimizing teacher feedback, considering and negotiating writing topic choice, engaging students in collaborative writing and assessment activities, getting them to use technological tools in writing learning, adapting teaching materials, cultivating students' writing motivational beliefs, incentivizing their participation in classroom activities, and relieving students' concerns about making errors. The reported frequencies of these motivational strategies indicate that some of them are more commonly used than others. The motivational strategy with the highest frequency is optimizing feedback ($n = 22$ teachers). According to the 22 teachers, they tried to optimize their feedback through different strategies, including: providing students with constructive and timely feedback, using individual feedback more than group (i.e., whole class) feedback, varying feedback content and focus, alleviating criticism in individual feedback, and referring to texts anonymously in group feedback.

With regard to the issue of considering and negotiating writing topic choice, the 17 teachers reporting using this motivational strategy said they select the topics matching students' interests and appropriate to their background knowledge, assign students easy writing tasks, engage them in choosing the topics they want to write about, or getting students to perform the one task as separate sub-tasks (i.e., planning, writing and revising). In the following interview, a female teacher is referring to one of these approaches in writing topic selection:

When students find topics irrelevant or not interesting, I try my best to personalize writing topic as per their interest. ... For example, if

TABLE 2 The teachers' perceptions of their students' writing de-motivation sources.

The source of writing de-motivation	No. of interviewees referring to it	Sample interview excerpt
Poor linguistic knowledge and writing ability	19	<i>There could be various factors causing students' lack of motivation to write in English, but I think the main factor is their poor English knowledge or their poor level in English essay writing, particularly in vocabulary. ... Many students usually avoid participating in classroom activities in order not to be criticized for their language level. (Teacher 1)</i>
Inappropriate writing topics	8	<i>Students lose motivation when they feel the writing topic is irrelevant or not interesting. ... They will not participate in the writing activities because they have no idea about the topic.... I also notice this case of losing motivation when students feel the topic is more difficult than expected; I mean it does not match their academic level. (Teacher 29)</i>
Ineffective teaching materials and techniques	7	<i>If textbooks are boring or teaching is monotonous, it will be difficult for students to stay motivated. ... Lack of writing motivation can be also caused by any brief feedback students get from their writing teachers. ... Students may struggle immensely if they have unclear feedback. (Teacher 10)</i>
Previous poor writing learning experiences	4	<i>Another cause of students' negative writing motivation is their poor level in writing during the pre-university stage... They did not learn how to write good English texts in schools. (Teacher 14)</i>
Cognitive nature of writing tasks	2	<i>Some students feel that writing in general is difficult, and unlike practicing other language skills, they feel writing takes a long time. (Teacher 32)</i>

TABLE 3 The teachers' reported motivational strategies.

The motivational strategy	No. of interviewees referring to it	Sample interview excerpt
Optimizing teacher feedback	22	<i>I try to foster a positive feedback culture to avoid students' sensitivity to criticism in essay comments. ... I tend to write a positive comment before the negative one, and change the areas of writing I praise. ... I also explain the reason for any criticism and say it does not reflect the students' level. ... If comments are given to the whole class, they are always without names. (Teacher 11)</i>
Considering and negotiating writing topic choice	17	<i>I try to choose the writing topic students like or have ideas about. ... Sometimes, I allow students to choose the writing topic. ... In the classroom, I also get them to perform writing task in separate stages. I mean they first do planning stage, and then the writing and revising stages. (Teacher 24)</i>
Engaging students in collaborative writing and assessment activities	13	<i>I usually use pair or group work writing tasks to create a cooperative learning environment so that students feel motivated and motivate each other, and alleviate anxiety for everyone. ... I believe encouraging students to participate in writing activities is the way to go. (Teacher 23)</i>
Getting students to use technological tools in writing learning	12	<i>I use apps like Grammarly in classrooms to help students notice their errors. But for lower-level students, I use Nearpod or Google Docs to encourage collaborative writing and discussion. (Teacher 2)</i>
Adapting teaching materials and techniques	10	<i>Even though we adhere to the prescribed textbook, I use additional and more engaging supplementary materials. ... I try to diversify the materials and use extra materials to enhance students' understanding and motivation. (Teacher 15)</i>
Cultivating students' writing motivational beliefs	9	<i>Teacher 33: At the very beginning of each writing course, I always try to show them the importance of writing to their future career and to proficiency in English. ... I explain that writing holds substantial relevance in real-world contexts.</i>
Incentivizing students' classroom participation	8	<i>I make efforts to help students with low writing de-motivation become more motivated in my classes. I assign marks for participation. (Teacher 1)</i>
Relieving students' concerns about making errors	3	<i>I try to push students to be motivated in their writing and not to worry about errors. ... I explain they are not held accountable for grammatical or stylistic errors. I always say to them: the more mistakes you make, the better writer you become. (Teacher 9)</i>

a writing unit discusses sports, I would make the task to be about a sport they like or wish to do it in real life ... This sort of personalizing writing tasks really encourages students to write. (Teacher 16)

Thirteen teachers mentioned engaging students in collaborative activities as a way for motivating them to write. Seven out of these 13 teachers mentioned using collaborative activities in the form of pair or group work writing tasks, while the other six teachers reported using peer assessment activities. It is noteworthy that the majority of the interviewees ($n = 27$) had a negative attitude toward using peer assessment. According to these teachers, peer assessment could have detrimental effects on students with low writing competence, sensitivity to peer criticism, or lack of seriousness. The following three teachers elaborated on this point as follows:

I do not prefer getting students to evaluate the work of their classmates. Students may feel very sensitive and embarrassed. Generally and typically, students exhibit reluctance in evaluating their peers' essays. (Teacher 3)

Unfortunately, my students' levels are not that good.... A few students can add to their classmates. ... So, I prefer to avoid getting them to correct their peers' errors. (Teacher 19)

I used to do peer evaluation several years ago, but I noticed that most students deliberately write nice comments and do not point out errors. (Teacher 21)

The narratives of the six teachers who reported using peer assessment activities indicate they implement them non-regularly and cautiously. Two of these teachers said they use these activities a few times a term, while the other four teachers said they use them conditionally, for instance after students know each other very well, or when they have a suitable writing competence level. Three teachers also said that preparing students for these activities is another complicated issue. The following interview excerpts show these cases:

I use peer assessment just to break the ice, and help students notice their errors.... But I use it only four weeks after the beginning of the writing course because peer revision is resisted when learners do not know each other. ... As time passes, students' shyness decreases. ... It always helps to use a rubric for these tasks. (Teacher 7)

I use peer evaluation if students in my class are of intermediate and higher levels only. In pre-intermediate classes, peer assessment normally causes much embarrassment for low-level students. ... Before getting students to participate in peer evaluation activities, I often establish clear expectations and guidelines and create a supportive and respectful classroom environment in order to overcome resistance. ... I then observe the interaction of students in each group to make sure everyone participates. I also assist students when necessary. (Teacher 16)

I help students have a positive peer evaluation experience by addressing their concerns, and by fostering a culture of respect and constructive criticism. ... I start with talking to students individually to find if they resist participating in peer evaluation activities. Once I understand their concerns, I can deal with them. I also try to help them understand the benefits of peer evaluation. ... To help students have a guided peer assessment activity, I provide guidance on giving

constructive feedback, and monitor their feedback participation. (Teacher 20)

The above interview excerpts suggest that implementing peer assessment in L2 writing courses requires some particular conditions leading to the desired motivational impact.

As for the role of technology in motivating students to write, the interviewed teachers were divided about this issue. The larger group of the teachers reported a negative attitude toward using technological tools in writing classes. For these teachers, using technological tools in writing classes is not beneficial and can cause unfavorable outcomes:

I believe technology affects students negatively. So, I do not allow them to use it in the classroom because if students were introduced to particular applications, essay plagiarism and auto-correction will increase. ... Students will learn writing better without technology. (Teacher 11)

I do not feel technology is important in writing classes. It's better to give the student the opportunity to try, make mistakes, and find the solution. But I only encourage my students to use grammar and spelling checking tools at home, and they usually like them. (Teacher 28)

The 12 teachers who mentioned making use of technology in writing classes to motivate students referred to using tools such as Grammarly, Nearpod or Google Nearpod, Google Docs or blogs. Overall, these teachers' answers indicate they do not make great use of technology in their writing instruction. Like their attitude toward technology use, the larger number of the teachers did not feel a dire need for adapting teaching materials for fostering students' motivation. These teachers viewed that students rarely get dissatisfied with the writing teaching materials used. Additionally, eight out of the 10 other teachers narrating making some kind of language teaching material adaptation said that it is only contingent upon noting dissatisfaction and de-motivation symptoms.

Only nine teachers talked about cultivating students' writing motivational beliefs. Their narratives showed they care about some of these motivational beliefs rather than others. Specifically, the nine teachers mainly referred to their attempts to cultivate students' motivation through highlighting the value of writing to their academic life and future careers, and, to a less extent, trying to alleviate students' negative writing attitude. Apart from this, no teacher talked about cultivating other motivation dimensions such as helping students regulate their emotions while writing, promoting their writing self-ability beliefs, or supporting them in setting achievable goals in writing courses. Eight teachers reported trying to motivate students by incentivizing their participation in classroom activities. These teachers' common strategy was to make a part of students' coursework marks dependent upon their active participation in such activities. Likewise, three teachers mentioned they try to encourage students to write by relieving their concerns about making written errors. For these teachers, relieving students' fears of error criticism could be one way for encouraging them to write.

Finally, it is worth noting that three teachers reported they do not feel obliged to care about students' writing motivation. In other words, these teachers believed that such motivating task is not a main part of their instructional roles. In the following interview excerpts, two of

them obviously stated they pay little attention to considering students' motivation in their classes:

The only way I tried in motivating students is choosing a topic that could match their interest or getting them to choose a topic they are interested in. (Teacher 4)

I do not consider students' opinions about my criticism or praise worthy of listening to. I only provide students with detailed written and corrective feedback. (Teacher 26)

Additionally, the interviews indicate the little attention the teachers pay to directly cultivating students' writing motivation. When asked about this issue, one female teacher, for instance, commented, "I do not really care about it if students remain de-motivated" (Teacher 7).

The strategies the teachers reported using do not reflect a wide range of writing motivation procedures when compared to the previous relevant taxonomies (e.g., Abdel Latif, 2021; Bruning and Horn, 2000), see section 2. For example, the teachers' reported motivational strategies do not have much in common with Bruning and Horn's (2000) guidelines for cultivating students' functional beliefs about writing, and creating a supportive learning atmosphere. These strategies do not either adequately cover Abdel Latif's (2021) proposed guidelines for meeting students' language needs, orchestrating peer assessment activities and nurturing students' writing motivational perceptions. The finding that some teachers in the Saudi university context are not interested in motivating their students may partially account for the noted limited conceptualizations of writing de-motivation (see 5.1). Therefore, more attention should be paid to raising teachers' awareness of the value of and effective ways for promoting students' writing motivation.

5.3 Students' writing competence and classroom size as potential correlates of the teachers' use of motivational strategies

The teachers' interview answers also helped in understanding their view on the role of students' writing competence and classroom size as potential factors influencing their use of motivational strategies in L2 writing classes. Regarding the role of students' writing competence, most teachers ($n = 24$) said it does not greatly influence their use of varied motivational strategies. For some of these teachers, competent student writers are normally motivated and therefore they only motivate low-level students who are de-motivated as a result of writing skill deficiencies. For another team of these teachers, motivational strategies remain unchanged when dealing with both high- and low-level writers.

Conversely, only nine teachers mentioned varying strategies for motivating students with different writing levels. The following two cases are typical of the narratives reported by the teachers in this group:

Good students usually struggle with uninteresting topics. ... I try to make the topics personal and ask them to search for background information to get their ideas flow. ... Poor students struggle with writing mechanics and basics and feel overloaded easily. So, I try to

ask them to use Grammarly to know their errors and edit their writing instead of feeling stressed about everything. (Teacher 12)

Students with good performance are always motivated. Many writing activities and tasks in the course are too easy for them. So, what they need is to work on more difficult tasks. For example, I get them to do more challenging writing activities during classes, leaving easier activities to poor students as they will feel more motivated when answering them correctly. (Teacher 30)

Collectively, the nine teachers mentioned motivating high-level students through changing writing topics and getting them to do more challenging writing activities or tasks, and they reported motivating low-level students through getting them to do easier tasks, supporting their learning, or guiding them to digital resources for completing their essays or revising them.

Regarding class size, all the teachers congruently viewed that it is a more influential factor in their use of writing motivation strategies. According to them, the fewer number of students are easier to motivate and to pay individualized attention to; for example:

Of course it is better to have a fewer students so that you have the time to focus on each student. So, a smaller group tends to be more manageable. In larger classes, it may be more challenging to provide individualized attention to make effective use of strategies such as peer collaboration and using technology for personalized feedback. In smaller classes, more individualized motivational support and one-on-one feedback can be provided. (Teacher 8)

Certainly, the number of students in the course can influence the solutions for related problems. In larger classes, group activities and collaborative writing can encourage participation and engagement. In smaller classes, more attention, individualized feedback and tailored instruction can be provided to deal with specific challenges related to writing motivation. (Teacher 19)

The above two interview parts suggest that even if the teachers are willing to exert considerable efforts in fostering students' writing motivation, the large number of students in one class can hinder their task. That is why minimizing students' numbers in English writing classes is key to helping the teachers in accomplishing their motivational task.

6 Discussion and conclusions

The present study aimed at exploring Saudi university EFL teachers' diagnosis of students' writing de-motivation symptoms (i.e., signs or indicators) and causes, and the strategies they use to motivate students to write and participate actively in classroom activities, i.e., their motivational strategies. Overall, we did not note tangible differences among the teacher groups at the five universities in these three dimensions; they were generally similar in their diagnosis of writing de-motivation indicators and causes, and also in their reported motivational strategies. The study uncovered important results about the teachers' writing motivation literacy and the efforts they allocate to get their student writers motivated. Compared to the previous few studies (Cheung, 2018; Lee and Lin, 2022; Mali, 2017; Rosina, 2017; Saranraj et al., 2014), using semi-structured interviews and framing

data collection and analysis based on a more relevant framework (Abdel Latif, 2021) have helped in revealing different insights into the realities of motivational strategies in L2 writing classes.

With regard to the teachers' conceptualizations of writing de-motivation, these were found to be rather limited as it turned out they lack a comprehensive awareness level of what it involves. Obviously, the teachers' conceptualizations fall only in the attitudinal and ability belief dimensions of writing de-motivation (see Abdel Latif, 2021). In other words, the de-motivation symptoms mentioned by the teachers concern students' negative attitudes toward writing, and avoidance behaviors, task procrastination, and writing block. The teachers also referred to poor linguistic knowledge and writing ability – an ability belief dimension – as a cause of students' de-motivation. Apart from attitudes and ability beliefs, the teachers mentioned no other ones related to the situational or goal orientation dimensions of students' de-motivation. While it is well-acknowledged that these signs may vary from one student and/or context to another, de-motivated writers normally experience attitudinal, situational, ability belief, and goal orientation symptoms – see Abdel Latif's (2021) definition in section 2. Thus, writing de-motivation is not merely an attitudinal-ability belief construct but it is a four-dimensional one. The teachers' limited conceptualization of writing de-/motivation could have resulted from a lack of awareness of its multiple aspects. It may have also been associated with the little attention some teachers paid to cultivating students' writing motivation (as indicated in subsection 5.2). The teachers' diagnosis of writing de-motivation causes is generally consistent with pertinent literature and previous research findings (e.g., Abdel Latif, 2021; Karaca and Inan, 2020; Pajares et al., 2007). It seems that writing de-motivation sources/causes occur consistently across different international educational settings.

Compared to previous research findings (e.g., Cheung, 2018; Lee and Lin, 2022; Rosina, 2017), the present study revealed a wider range of teacher motivational strategies in Saudi university English writing classes. However, the strategies the teachers mentioned using relate only to some dimensions in the motivational strategy frameworks proposed by Bruning and Horn (2000) and Abdel Latif (2021). Specifically, the teachers' motivational strategies align adequately with the proposed guidelines of using appropriate teaching materials and writing tasks, and optimizing teacher feedback, but they align partially with the guidelines of developing a positive learning environment, cultivating students' writing motivational perceptions and creating a motivating learning atmosphere via using technological tools in instruction. Meeting students' language and writing performance needs and orchestrating peer assessment activities seem to be almost neglected as the teachers mentioned a very few motivational strategies pertaining to these two guidelines. Consistent with previous language motivation research findings (e.g., Cheng and Dörnyei, 2007; Dörnyei, 1998; Guilleaumeaux, 2013), the present results indicate that teachers are likely to be unable to use motivational strategies effectively in writing classes with a large number of students. Meanwhile, the study has not provided conclusive evidence for the interaction between teachers' use of motivational strategies and students' writing competence levels.

In light of the present results, it is concluded that Saudi university teachers' motivational practices in English writing classes are yet to be enhanced. Such enhancement requires three steps. First, teachers' writing motivation literacy or awareness needs be fostered to help them understand writing motivation dimensions and know how to

motivate students properly. Second, there is also a need for activating the use of motivational strategies in writing classes. Both steps could be accomplished through teacher education programmes and in-service training workshops. Besides, teacher educators and textbook writers have the potential to play an important role in drawing teachers' attention to relevant issues. For example, pre-service and in-service teacher educators may pay due attention to raising teachers' awareness of writing de-motivation symptoms and causes, and to fostering their consciousness of how to help de-motivated students. Teaching supervisors could also draw teachers' attention to some guidelines for optimizing their practices to meet students' writing motivational needs. In addition, teacher guides could include some instructional scenarios related to diagnosing students' writing de-motivation symptoms and alleviating them. Congruent with the present results emphasizing the central role of class size in enabling teachers' use of motivational strategies, there is also a need for minimizing the number of students in English writing classes at Saudi universities. With appropriate writing class size, teachers' task in getting students motivated will be easier.

The scant research on writing teachers' motivational strategies calls for addressing multiple dimensions in this area. The realities of motivational strategy use need to be investigated in different international L2 writing learning settings. It is also important to profile a larger number of factors potentially influencing writing teachers' use of motivational strategies, including students' gender, educational stage and cultural context. Another issue worth investigating in future studies is how teachers' writing motivation literacy may impact the ways they motivate students in writing classes. Future relevant studies can also draw upon other data sources such as classroom observation or combine it with semi-structured interviews or questionnaires. Such future research could help in disseminating a more effective motivation culture in L2 writing learning environments.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

This study was approved by The Ethics Committee of the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University. The study was conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MA: Conceptualization, Formal analysis, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. TALg: Data curation, Formal analysis, Investigation,

Methodology, Validation, Writing – original draft, Writing – review & editing. TAs: Data curation, Formal analysis, Funding acquisition, Investigation, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number: IMSIU-DDRSP2504).

References

- Abdel Latif, M. M. M. (2011). What do we know and what do we need to know about Arab gulf EFL/ESL students' writing? *TESOL Arabia Perspec.* 18, 6–14.
- Abdel Latif, M. M. M. (2015). Sources of L2 writing apprehension: a study of Egyptian university students. *J. Res. Read.* 38, 194–212. doi: 10.1111/j.1467-9817.2012.01549.x
- Abdel Latif, M. M. M. (2019a). Unresolved issues in defining and assessing writing motivational constructs: a review of conceptualization and measurement perspectives. *Assess. Writ.* 42:100417. doi: 10.1016/j.asw.2019.100417
- Abdel Latif, M. M. M. (2019b). Helping L2 students overcome negative writing affect. *Writ. Pedagog.* 11, 151–163. doi: 10.1558/wap.38569
- Abdel Latif, M. M. M. (2021). Writing motivation research, measurement, and pedagogy. London: Routledge.
- Abdel Latif, M. M. M., Alghizzi, T., and Alshahrani, T. (2024). The impact of genre-based instruction on Saudi university students' 39 English writing performance and motivation: a mixed-method study. *Front. Psychol.* 15, 1–9. doi: 10.3389/fpsyg.2024.1387009
- Abdel Latif, M. M. M., Alsahil, A., and Alsuhailani, Z. (2024). Profiling EFL writing teachers' feedback provision practices and activity uses in Saudi universities. *Human. Soc. Sci. Commun.* 11, 1–10. doi: 10.1057/s41599-024-03639-5
- Abdel Latif, M. M. M., Alsuhailani, Z., and Alsahil, A. (2024). Matches and mismatches between Saudi university students' English writing feedback preferences and teachers' practices. *Assess. Writ.* 43:100438. doi: 10.1016/j.asw.2019.100438
- Alsahil, A., Abdel Latif, M. M. M., and Alsuhailani, Z. (2024). Exploring EFL learner engagement with different teacher feedback modes. *Comput. Assist. Lang. Learn.*, 1–28. doi: 10.1080/09588221.2024.2374785
- Atay, D., and Kurt, G. (2006). Prospective teachers and L2 writing anxiety. *Asian EFL J.* 8, 100–118.
- Beshir, H. (2017). Assessing English language teachers' awareness and implementation of motivational strategies in EFL class: Chilalo and Tokuma high schools in Oromia region Arsi zone in focus, Ethiopia. *Int. J. Multicult. Multirelig. Underst.* 5, 203–210. doi: 10.18415/ijmmu.v5i3.358
- Bruning, R., and Horn, C. (2000). Developing motivation to write. *Educ. Psychol.* 35, 25–37. doi: 10.1207/S15326985EP3501_4
- Camacho, A., Alves, R. A., and Boscolo, P. (2021). Writing motivation in school: a systematic review of empirical research in the early twenty-first century. *Educ. Psychol. Rev.* 33, 213–247. doi: 10.1007/s10648-020-09530-4
- Cheng, H. F., and Dörnyei, Z. (2007). The use of motivational strategies in language instruction: the case of EFL teaching in Taiwan. *Innov. Lang. Learn. Teach.* 1, 153–174. doi: 10.2167/illt048.0
- Cheung, Y. L. (2018). The effects of writing instructors' motivational strategies on student motivation. *Aust. J. Teach. Educ.* 43, 55–73.
- Cruz Cordero, T., Wilson, J., Myers, M. C., Palermo, C., Eacker, H., Potter, A., et al. (2023). Writing motivation and ability profiles and transition during a technology-based writing intervention. *Front. Psychol.* 14:1196274. doi: 10.3389/fpsyg.2023.1196274
- Dörnyei, Z. (1998). Demotivation in foreign language learning. Paper presented at the TESOL '98 congress, Seattle, WA, March.
- Dörnyei, Z. (2000). Motivation in action: towards a process-oriented conceptualization of student motivation. *Br. J. Educ. Psychol.* 70, 519–538.
- Dörnyei, Z. (2001). Motivational strategies in the foreign language classroom. Cambridge: Cambridge University Press.
- Dörnyei, Z., and Csizér, K. (1998). Ten commandments for motivating language learners: results of an empirical study. *Lang. Teach. Res.* 2, 203–229. doi: 10.1177/136216889800200303
- Drew, I., and Sorheim, B. (2009). English teaching strategies. Oslo, Norway: Det Norske Samlaget.
- Ferris, R. D. (2002). Treatment of error in second language writing. Ann Arbor: University of Michigan Press.
- Guilloteaux, M. J. (2013). Motivational strategies for the language classroom: perceptions of Korean secondary school English teachers. *System* 41, 3–14. doi: 10.1016/j.system.2012.12.002
- Guilloteaux, M. J., and Dörnyei, Z. (2008). Motivating language learners: a classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Q.* 42, 55–77. doi: 10.1002/j.1545-7249.2008.tb00207.x
- Karaca, M., and Inan, S. (2020). A measure of possible sources of demotivation in L2 writing: a scale development and validation study. *Assess. Writ.* 43:100438. doi: 10.1016/j.asw.2019.100438
- Karlen, Y., and Compagnoni, M. (2017). Implicit theory of writing ability: relationship to metacognitive strategy knowledge and strategy use in academic writing. *Psychol. Learn. Teach.* 16, 47–63. doi: 10.1177/1475725716682887
- Kim, Y. K., and Kim, T. Y. (2024). Motivational languaging intervention for L2 learners: the differential effect of individual and group writing. *Lang. Teach. Res.* 28, 894–912. doi: 10.1177/13621688211011498
- Lee, T. S., and Lin, L. H. (2022). Motivational strategies adopted in postgraduate English academic writing courses: perspectives of doctoral students. *SAGE Open* 12, 1–11. doi: 10.1177/21582440221138259
- Limpo, T., and Alves, R. (2017). Relating beliefs in writing skill malleability to writing performance: the mediating role of achievement goals and self-efficacy. *J. Writ. Res.* 9, 97–125. doi: 10.17239/jowr-2017.09.02.01
- Lo, J., and Hyland, F. (2007). Enhancing students' engagement and motivation in writing: the case of primary students in Hong Kong. *J. Second. Lang. Writ.* 16, 219–237. doi: 10.1016/j.jslw.2007.06.002
- Lodico, M. G., Spaulding, D. T., and Voegtli, K. H. (2006). Methods in educational research from theory to practice. San Francisco: Jossey-Bass.
- Mali, Y. C. G. (2017). Motivational teaching strategies in Indonesian EFL writing classrooms. *I 17*, 60–85. doi: 10.24167/celt.v17i1.1139
- Merriam, S. B. (1998). Qualitative research and case study applications in education. San Francisco: Jossey-Bass.
- Pajares, F., Johnson, J., and Usher, E. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Res. Teach. Engl.* 42, 104–120.
- Reeves, L. L. (1997). Minimizing Iwriting apprehension in the learner centered classroom. *Engl. J.* 86, 38–45. doi: 10.2307/820367
- Rosina, Y. (2017). Motivational strategies with particular emphasis on writing when teaching English in Norwegian schools. Norway: University of Stavanger.
- Saranraj, L., Zafar, S., and Khan, Z. A. (2014). Teachers' use of motivational strategies in tertiary level ESL classrooms. Proceedings of the International Conference on Trends and Innovations in Language Teaching, 462–466.
- Shen, B., and Bai, B. (2024). Enhancing Chinese university students' writing performance and self-regulated learning (SRL) writing strategy use through a strategy-based intervention. *System* 122:103249. doi: 10.1016/j.system.2024.103249

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Teng, L. S., Yuan, R. E., and Sun, P. P. (2020). A mixed-methods approach to investigating motivational regulation strategies and writing proficiency in English as a foreign language contexts. *System* 88, 1–12. doi: 10.1016/j.system.2019.102182
- Troia, G., Shankland, R., and Wolbers, K. (2012). Motivation research in writing: theoretical and empirical considerations. *Read. Writ. Q.* 28, 5–28. doi: 10.1080/10573569.2012.632729
- Tsao, J., Tseng, W., and Wang, W. (2017). The effects of writing anxiety and motivation on EFL college students' self-evaluative judgments of corrective feedback. *Psychol. Rep.* 120, 219–241. doi: 10.1177/0033294116687123
- Waddington, J. (2018). Teacher understanding and implementation of motivational strategies in ELT. *ELT J.* 72, 162–174. doi: 10.1093/elt/ccx044
- Walker, B. (2003). The cultivation of student self-efficacy in reading and writing. *Read. Write. Q.* 19, 287–302. doi: 10.1080/10573560308217
- Yu, S., Jiang, L., and Zhou, N. (2020). Investigating what feedback practices contribute to students' writing motivation and engagement in Chinese EFL context: a large scale study. *Assess. Writ.* 44:100451. doi: 10.1016/j.asw.2020.100451
- Zarrinabadi, N., Mahmoudi-Gahrouei, V., and Mohammadzadeh Mohammadabadi, A. (2023). Your writing could have been better: examining the effects of upward and downward counterfactual communication on the motivational aspects of L2 writing. *Assess. Writ.* 56:100714. doi: 10.1016/j.asw.2023.100714

Appendix 1

Guiding interview questions

1. I would like to start this interview by getting some information about your language and writing teaching experiences. Could you please tell me about your writing instruction and the writing courses and students you have taught?
2. In each writing class, there are normally some students who do not feel motivated to write or do not like to participate in classroom activities. When and how do you usually notice students with inadequate writing motivation? Please explain in detail the behaviors of such students.
3. In your opinion, what have caused these students not to feel motivated to write in English? Please explain in detail any potential related causes you have noted in your English writing classes.
4. Do you usually try to help students with low writing motivation in your classes become more motivated? If so, please explain in detail the things you do or the strategies you use for this purpose.
5. To what extent do students in your English writing classes feel happy or satisfied with the textbooks or instructional techniques you use? If they are not happy with them, how do you deal with this problem? Please explain in detail.
6. How many essay topics do you normally assign your students each term in your writing courses? What are the students' reactions to the number and types of the essay topics assigned? How do you try to overcome any problems related to the students' reactions to the number of essay assignments or essay topics? Please explain in detail.
7. Some students with good or poor writing performance alike may have a lack of writing motivation. Have you noted this case in your English writing classes? Generally, how do you try to help each type of students to become motivated? Please explain in detail.
8. Do you make use of technology for fostering your students' motivation to write and to get them engaged in English writing classroom activities? If so, which technological aids or tools do you use in your English writing classes? And what are students' reactions to them? Please explain in detail.
9. When commenting on students' English essays, do some of them feel annoyed with your criticism or even bored with your repeated praise? Is this the same when giving oral or written, to an individual student or a group of students- or to students with different performance improvement levels? How do you usually try to consider students' potential sensitivity to or boredom with your comments on their essays? Please narrate your related experiences.
10. In your English writing classes, do you get students to evaluate their classmates' essays? In case you have noticed some students resist participating in this type of evaluation activities, how do you usually try to help these students have a positive peer evaluation experience? Please explain your related experiences in detail.
11. When dealing with all the above mentioned issues in your writing classes, do your solutions of any potential writing motivation-related problems differ depending on:
 - The number of students in the course.
 - The students' academic level or progress.

Please explain in detail.



OPEN ACCESS

EDITED BY

Xinghua Liu,
Shanghai Jiao Tong University, China

REVIEWED BY

Joseline Santos,
Bulacan State University, Philippines
Vivek Bhardwaj,
Manipal University Jaipur, India

*CORRESPONDENCE

Min Zhang
✉ xaiu13176@xaiu.edu.cn

RECEIVED 04 June 2025

ACCEPTED 25 August 2025

PUBLISHED 01 October 2025

CITATION

Zhang M (2025) Optimizing academic engagement and mental health through AI: an experimental study on LLM integration in higher education. *Front. Psychol.* 16:1641212. doi: 10.3389/fpsyg.2025.1641212

COPYRIGHT

© 2025 Zhang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Optimizing academic engagement and mental health through AI: an experimental study on LLM integration in higher education

Min Zhang *

College of Humanities and Arts, Xi'an International University, Xi'an Shaanxi, China

Background: In alignment with UNESCO's Sustainable Development Goal 4 (SDG4), which advocates for inclusive and equitable quality education, the integration of Artificial Intelligence tools—particularly Large Language Models (LLMs)—presents promising opportunities for transforming higher education. Despite this potential, empirical research remains scarce regarding the effects of LLM use on students' academic performance, mental well-being, and engagement, especially across different modes of implementation.

Objective: This experimental study investigated whether a guided, pedagogically grounded use of LLMs enhances students' academic writing quality, perceived mental health, and academic engagement more effectively than either unguided use or no exposure to LLMs. The study contributes to UNESCO's "Futures of Education" vision by exploring how structured AI use may foster more inclusive and empowering learning environments.

Method: A total of 246 undergraduate students were randomly assigned to one of three conditions: guided LLM use, unguided LLM use, or a control group with no LLM access. Participants completed a critical writing task and standardized instruments measuring academic engagement and mental well-being. Prior academic achievement was controlled for, and writing quality was assessed using Grammarly for Education.

Results: Students in the guided LLM condition achieved significantly higher scores in writing quality and academic engagement compared to the control group, with large and moderate effect sizes, respectively. Modest improvements in mental health indicators were also observed. By contrast, unguided use yielded moderate gains in writing quality but did not produce significant effects on engagement or well-being.

Conclusion: The findings highlight the critical role of intentional instructional design in the educational integration of AI tools. Structured guidance not only optimizes academic outcomes but also supports students' wellbeing and inclusion. This study offers empirical evidence to inform ongoing debates on how digital innovation can contribute to reducing educational disparities and advancing equitable learning in the post-pandemic era.

KEYWORDS

Large Language Models, academic writing, mental health, student engagement, higher education, guided instruction, educational technology, AI-assisted learning

1 Introduction

The accelerated integration of Artificial Intelligence (AI) into higher education is reshaping the academic landscape at an unprecedented pace. In particular, the emergence of Large Language Models (LLMs), such as OpenAI's ChatGPT, has generated both enthusiasm and apprehension among educators and policymakers. While some institutions have embraced these technologies as tools to enhance personalization, accessibility, and innovation in teaching, others have expressed concern about academic integrity, student dependency, and the erosion of critical thinking. The current moment thus presents a pivotal opportunity—and challenge—for universities to evaluate the pedagogical value of LLMs and their broader impact on student learning (Sharma et al., 2025).

The urgency of this evaluation is underscored by the widespread and rapid adoption of LLMs in academic contexts. For instance, recent headlines such as “More than half of UK undergraduates say they use AI to help with essays” (Adams, 2024) reflect a shift in student practices that institutions are still struggling to regulate or harness effectively (Fritz et al., 2024). Despite this proliferation, empirical evidence remains limited, especially regarding how different modalities of LLM implementation—guided versus unguided use—affect students' academic performance, mental wellbeing, and engagement. Given the scale and speed of adoption, addressing this gap has become an urgent priority for educators and researchers alike.

In this context, international policy frameworks such as UNESCO's Sustainable Development Goal 4 (SDG4), which promotes inclusive, equitable, and quality education, and the “Futures of Education” initiative offer critical guidance. The latter calls for reimagining how knowledge is produced, valued, and shared, with a strong emphasis on human-centered, ethically grounded digital innovation. This vision aligns closely with the need to understand how emerging technologies like LLMs can support not only academic excellence, but also psychological wellbeing and inclusive engagement among students.

Integrating AI into university education is not merely a matter of technological adaptation; it compels a re-examination of core pedagogical processes. Academic writing, for example, remains a central yet often stressful academic demand, both difficult to master and to assess objectively (Ayeni et al., 2024). Simultaneously, student mental health has emerged as a pressing concern in higher education, particularly within competitive and international environments (Molodynski et al., 2021). Academic engagement—the emotional, cognitive, and behavioral investment in learning—is equally critical, yet sensitive to instructional design and motivation (Lin, 2024).

Although interest in educational applications of LLMs is growing (Ng et al., 2024), including recent efforts to synthesize their contributions to personalized learning (Sharma et al., 2025), few studies have experimentally assessed their impact on these three domains within controlled settings (Jungheer, 2023). Furthermore, how these tools are introduced—whether with structured guidance or left to student discretion—may significantly influence their effectiveness and students' emotional and cognitive responses to academic tasks (Chang, 2024).

The present study addresses this research gap by experimentally examining the effects of guided versus unguided use of an LLM on undergraduate students' academic writing quality, perceived mental health, and academic engagement. Conducted in an international university in China, the study employed a standardized writing task and randomized group assignment (guided use, unguided use, control) to determine whether structured integration enhances learning outcomes and wellbeing. The results aim to inform evidence-based, ethical practices for AI integration in higher education and contribute to global discussions on how digital tools can advance more inclusive, resilient, and human-centered academic environments.

2 Theoretical framework and empirical background

2.1 LLMs in higher education

LLMs, such as GPT-4, are increasingly present in higher education as tools to assist with language production, research synthesis, and academic writing (Lu et al., 2024). Their growing use among university students has sparked institutional interest in understanding how these tools influence learning outcomes. However, emerging evidence suggests that the pedagogical value of LLMs depends less on their availability than on the instructional design that accompanies their use (Robledo et al., 2024).

A useful framework for analyzing the educational integration of technology is the Technological Pedagogical Content Knowledge (TPACK) model developed by Mishra and Koehler (2006). TPACK posits that effective technology-enhanced instruction requires the intersection of three types of knowledge: disciplinary content, pedagogical strategies, and technological tools. In this model, the mere introduction of digital resources does not guarantee meaningful learning. Rather, it is the thoughtful alignment of those tools with pedagogical goals and disciplinary content that fosters deep understanding and transferable skills.

This framework is particularly relevant to the use of LLMs. In a guided implementation, students receive explicit instructions on how to use the model to support key aspects of academic writing—such as developing argument structure, paraphrasing source material, or revising according to disciplinary conventions (Yan et al., 2024). This structured use of the tool reflects the TPACK ideal: technology embedded within a coherent pedagogical plan.

By contrast, unguided use of LLMs lacks this intentional alignment. Although students may independently explore the tool's capabilities, they do so without pedagogical framing, which may result in inconsistent outcomes. Unguided users might underuse the tool, rely on it uncritically, or fail to recognize its limitations (Wang, 2022). Finally, students in the control group, with no access to LLMs, must rely entirely on their prior writing skills and internal resources. While this condition mirrors traditional academic expectations, it may pose additional cognitive and emotional challenges for students with lower confidence or weaker academic preparation (Ayeni et al., 2024). Building on this theoretical foundation, the present study investigates

how different instructional approaches to LLM use—guided, unguided, or absent—affect academic writing, mental well-being, and engagement. The TPACK framework supports the hypothesis that pedagogically framed LLM use will yield superior outcomes across all domains.

3 Academic writing quality

Academic writing is a core component of higher education, particularly in the humanities and social sciences. It demands clarity of argument, mastery of disciplinary conventions, and grammatical precision—competencies that students often struggle to develop and instructors find difficult to evaluate objectively (Ayeni et al., 2024). Studies have shown that structured instructional approaches, such as modeling, scaffolding, and feedback, consistently improve students' writing skills (De La Paz, 2005). LLMs offer a new form of writing support, assisting students in generating ideas, organizing content, and refining their language. Early findings suggest that students who use these tools during the planning and revision phases may produce more coherent and technically accurate texts (Lee, 2023). However, the benefits of LLMs are not automatic. Their effectiveness hinges on how they are introduced and used in educational contexts.

From a TPACK perspective, guided LLM use can enhance academic writing by aligning the tool's features with pedagogical goals. Instructors may, for instance, teach students how to use the model to outline arguments or critically revise text while warning against uncritical copying or overreliance. This structured integration supports metacognitive engagement and allows students to internalize academic writing conventions. In contrast, students in the unguided condition may fail to use the tool optimally. Without pedagogical framing, they might use it only superficially—for grammar correction or idea generation—without fully engaging with the writing process. Additionally, they may be more prone to accept AI-generated suggestions uncritically, leading to errors in reasoning, style, or source use (Wang, 2022).

For students in the control condition, the writing task requires managing all stages of composition without external digital support. While this reflects a traditional academic scenario, it may impose greater cognitive demands and limit writing quality, especially for students lacking confidence or fluency in academic writing (Ayeni et al., 2024). Based on this reasoning, the study hypothesizes that students in the guided LLM condition will produce significantly higher-quality academic writing than those in the unguided and control groups, respectively. These differences are theoretically grounded in the TPACK framework and supported by prior research on instructional scaffolding and technology-mediated writing support.

4 Perceived mental health

University students' mental health has become a central concern in global higher education, with consistently high levels of anxiety, stress, and emotional exhaustion reported across diverse national contexts (Granieri et al., 2021). These issues are

particularly salient in competitive academic environments, where cognitive demands are high and support structures often limited. Academic writing, in particular, is a cognitively and emotionally taxing task that may exacerbate stress, especially in the absence of timely guidance or feedback.

The Cognitive Load Theory (Sweller, 1988) provides a relevant framework for understanding how instructional conditions affect students' mental wellbeing. According to this theory, cognitive performance is shaped by the interplay of three types of load: intrinsic (task complexity), extraneous (inefficient instructional design), and germane (learning-related processing). Poorly structured tasks tend to increase extraneous load, consuming cognitive resources and contributing to frustration or emotional fatigue (Li et al., 2020).

LLMs, when properly embedded in instruction, can help reduce extraneous cognitive load by automating lower-level processes such as sentence formulation, grammar correction, or even idea generation. However, this benefit is not automatic. Students need pedagogical guidance to understand how to use the tool effectively and ethically, and how to interpret or revise its suggestions. Without such framing, students may misuse the tool, become overwhelmed by its outputs, or develop dependency without comprehension (Park and Ahn, 2024).

In the guided condition, students receive structured instructions on how to use the LLM strategically during the writing process—e.g., to plan text sections, refine transitions, or paraphrase while maintaining academic integrity. This structure is expected to reduce cognitive overload and enhance students' sense of control, which may, in turn, support emotional regulation and perceived well-being. In contrast, the unguided group accesses the tool without clear direction. While they may benefit from its features, they also face the burden of interpreting outputs and deciding when and how to use them. This may increase cognitive load rather than reduce it, particularly for students unfamiliar with AI tools or lacking academic writing experience. Consequently, their perceived mental health may remain unchanged or even be negatively affected.

Finally, students in the control group, without access to any external tool or guidance, must complete the writing task using only their own cognitive and emotional resources. While this mirrors traditional academic practice, it may result in heightened task-related anxiety or emotional exhaustion, especially under time constraints or pressure to perform.

Based on this framework, the present study hypothesizes that students in the guided LLM condition will report significantly better mental well-being than those in the control group, with the unguided group expected to fall somewhere in between. This hypothesis reflects the assumption that instructionally structured technology use, rather than mere access, is the key to supporting psychological outcomes in academic settings.

5 Academic engagement

Academic engagement is a multidimensional construct encompassing students' behavioral, emotional, and cognitive investment in learning activities (Fredricks et al., 2004). High levels of engagement have been associated with greater academic

achievement, persistence, and satisfaction, particularly in university settings where autonomy and self-regulation are central to success (Wang, 2022). However, engagement is also sensitive to fluctuations in motivation, task design, and perceived support from instructors or institutional structures (Lin, 2024).

A useful framework for understanding the mechanisms that foster engagement is the Self-Determination Theory (SDT), proposed by Deci and Ryan (2000, Ryan and Deci, 2000). According to SDT, engagement flourishes when learners experience the fulfillment of three basic psychological needs: competence (feeling effective), autonomy (feeling self-directed), and relatedness (feeling connected and supported). Instructional strategies that enhance these dimensions are more likely to result in sustained engagement and intrinsic motivation.

In this context, the use of LLMs has the potential to support academic engagement—but only if implemented thoughtfully. In the guided condition, students receive clear instructions on how to use the tool to improve their writing in ways that foster self-efficacy and control. For example, they may learn to use the model to test different formulations, organize their ideas more efficiently, or revise their text in response to feedback. This structured support not only enhances competence, but also promotes autonomy, as students gain agency in managing complex academic tasks. In contrast, students in the unguided condition are left to navigate the LLM independently. While some may explore the tool productively, others may feel uncertain about how to use it effectively or ethically. This ambiguity can hinder perceived competence and reduce the motivational benefits typically associated with technology-enhanced learning. Without explicit pedagogical framing, LLM use may become a passive or confusing experience, diminishing its capacity to support sustained engagement.

Finally, students in the control group engage in the task without any digital support. Although this may reflect a traditional educational scenario, it offers limited opportunities to enhance autonomy or competence through external scaffolding. For some students, especially those with lower academic confidence, this condition may result in disengagement or surface-level effort. Building on Self-Determination Theory and recent findings on digital learning environments (Wang, 2022), the present study hypothesizes that students in the guided LLM condition will report the highest levels of academic engagement, followed by those in the unguided condition, with the control group expected to exhibit the lowest levels. This hierarchy reflects the assumption that pedagogically structured AI use can enhance both motivation and investment in academic tasks, provided that it supports students' psychological needs.

Taken together, the theoretical and empirical perspectives reviewed above suggest that the impact of AI tools in higher education depends not merely on access to the technology, but critically on how that technology is pedagogically framed and operationalized. While guided use of LLMs has the potential to support students' writing development, reduce extraneous cognitive load, and foster meaningful engagement, unguided use may result in uneven or superficial outcomes. Meanwhile, students who receive no digital support may face greater academic pressure and cognitive effort, particularly when completing complex tasks under time constraints.

To examine these assumptions, the present study adopts an experimental design comparing three conditions: guided LLM use, unguided LLM use, and a control group without access to LLMs. The outcomes under investigation—academic writing quality, perceived mental wellbeing, and academic engagement—were selected because they represent core dimensions of student success and are theoretically linked to instructional design and technological integration. Building on the reviewed literature, it is hypothesized that students in the guided LLM condition will outperform their peers across all three variables, followed by those in the unguided condition, with the control group expected to report the lowest levels of performance and well-being. This hypothesis reflects the view that it is not the technology itself, but rather the pedagogical structuring of its use, that determines its educational value.

5.1 Hypotheses

- H1: Students in the guided LLM use condition will demonstrate significantly higher academic writing quality than those in the unguided LLM use and control groups.
- H2: Students in the guided LLM use condition will report significantly higher levels of perceived mental health compared to students in the control group.
- H3: Students in the guided LLM use condition will exhibit significantly greater academic engagement than those in the control group.
- H4: Students in the unguided LLM use condition will demonstrate intermediate levels of academic writing quality, perceived mental health, and engagement, higher than those in the control group but lower than those in the guided use group.

6 Method

6.1 Transparency and openness

In this experimental study, we report how the sample size was determined and all inclusion criteria, manipulations, and outcome measures. All anonymized data are available via the Open Science Framework (https://osf.io/htejmf/view_only=54624dbd9f11467ea26242bae037e713). The data were analyzed using SPSS, version 29. No data were collected after the data analysis began. This study was not preregistered.

6.2 Participants

An a priori power analysis was conducted using G*Power (Faul et al., 2009) to determine the required sample size for a one-way ANCOVA with three groups and one covariate. Setting the alpha level at 0.05, power at 0.80, and anticipating a small to medium effect size ($f = 0.20$), the estimated minimum sample size was $N = 246$. The final sample consisted of two hundred and eighty eight undergraduate students enrolled in humanities and arts programs at an international university in China. Instructors

from four elective courses were initially contacted via internal mailing lists distributed by the College of Humanities and Arts and were invited to authorize data collection during one of their scheduled sessions. Once instructor consent was obtained, students were approached in person during class and invited to participate. Participants were recruited using a convenience sampling strategy, and participation was strictly voluntary. Students were informed that they could decline or withdraw at any point without penalty. No academic credit, compensation, or incentive was offered. Of the approximately three hundred and twenty five students approached across the four courses, two hundred and eighty eight undergraduate (88.6%) agreed to participate and completed all study components. The final sample included one hundred and sixty eight male students (58.3%) and one hundred and twenty female students (41.7%), ranging in age from 18 to 22 years ($M = 19.88$, $SD = 1.50$). All participants completed the writing task and self-report measures under supervised classroom conditions.

6.3 Ethical approval and informed consent

This study was reviewed and approved by the Institutional Review Board (IRB) of the Xi'an International University following the Declaration of Helsinki. Ethical approval was granted on [January 15th, 2024], under the reference number [approval ID, IRB/24/072-HUMARTS]. Before participation, all students received an information sheet outlining the purpose of the study, the nature of the tasks, the voluntary nature of their participation, and their right to withdraw at any point without penalty. Written informed consent was obtained from all participants. The consent form emphasized that participation was anonymous, data would be kept confidential, and results would be used solely for academic research. Participants were also informed that using the LLM was part of an experimental educational intervention and that their course grades would not be affected by their responses or participation.

6.4 Procedure

Instructors from four elective undergraduate courses in the humanities and arts were first contacted via internal mailing lists distributed by the College of Humanities and Arts. After receiving their consent to conduct the study during scheduled class time, students were invited in person to participate. The study was introduced at the beginning of the session, and all students were informed that participation was entirely voluntary and that they could withdraw at any time without consequences. Students who agreed to participate provided informed consent and completed the study during a supervised class session. Participants were randomly assigned to one of three experimental conditions: guided LLM use, unguided LLM use, or control. Random assignment was conducted at the individual level within each classroom using a pre-generated randomization list.

All participants were asked to complete the same academic writing task: a critical essay on the topic “The impact of

globalization on contemporary culture”, to be written in 45 m using a computer. Participants in the two experimental conditions used OpenAI’s GPT-4, accessed via a monitored institutional interface. No alternative platforms or personal devices were permitted. All students interacted with the same LLM under identical interface conditions. On average, participants in the experimental conditions spent between 25 and 35 m actively interacting with the LLM during the task. In the control condition, students received the following prompt: “Write a critical essay on the impact of globalization, using the provided readings. Structure your argument and support it with specific examples.” No access to LLMs or external writing tools was provided. In the unguided LLM use condition, students were given access to GPT-4 and instructed: “You may use the language model (LLM) in any way you find useful to complete your essay.” No additional instructions, training, or support were provided.

In the guided LLM use condition, participants received the following prompt: “Write a critical essay on the impact of globalization. Use the language model (LLM) to help you generate ideas, organize your arguments, and improve clarity. You may use it to explore different perspectives, revise paragraphs, or paraphrase content. Ensure that your essay reflects critical thinking, coherence, and academic style.”

Before beginning the writing task, this group received a brief 10-m in-class orientation delivered by the course instructor, based on a script prepared by the research team. The orientation covered three key elements: how to formulate effective prompts, how to evaluate AI-generated outputs critically, and how to use the tool ethically in academic contexts. After completing the writing task, participants responded to standardized self-report measures assessing perceived mental health, academic engagement, and a short demographic questionnaire. All responses were submitted digitally and anonymized prior to analysis. No pilot study was conducted prior to the implementation of the experiment.

6.5 Instruments

All instructions, writing prompts, the manipulation check, and the self-report measures—except for one—were administered in English, in accordance with the instructional language of the international university where the study took place. The only exception was the Utrecht Work Engagement Scale – Student version (UWES-S), which was administered in its validated Chinese version due to its demonstrated psychometric reliability in Chinese undergraduate populations.

6.6 Manipulation check

A manipulation check was administered immediately after the writing task to verify participants’ adherence to their assigned intervention condition. Participants responded to two closed-ended questions: (1) “Did you use the language model (LLM) while completing the essay?” (Yes/No), and (2) “Were you instructed on how to use the LLM?” (Yes/No). These items allowed the researchers to determine whether participants in the experimental

groups used the LLM as intended, and whether participants in the control group refrained from doing so.

Only participants whose responses were fully consistent with their assigned condition were retained for the main analyses. Specifically, inclusion criteria required that participants in the guided condition reported using the LLM with instructions, those in the unguided condition reported using the LLM without instructions, and those in the control condition reported not using the LLM. Participants who did not meet these criteria were excluded from the final dataset. As a result, the final sample included two hundred and forty six participants who successfully passed the manipulation check and were eligible for analysis.

6.7 Text quality

The quality of participants' academic writing was assessed using the Grammarly for Education platform (Grammarly, Inc., 2024). After completing the essay, the experimenter uploaded each text under standardized conditions. Grammarly automatically generated a Performance Score, ranging from 0 to 100, which served as the primary indicator of overall text quality. This composite score reflects the extent to which the writing adheres to grammatical norms, clarity, and effective communication, and it can be improved by addressing the platform's suggested revisions. In addition to the performance score, Grammarly provides detailed linguistic metrics, including word count, average word and sentence length, readability score (based on the Flesch scale) (Flesch, 1948), and vocabulary diversity (e.g., proportion of unique and rare words). These secondary indicators were reviewed to contextualize writing complexity and stylistic variation, though only the Performance Score was used in the statistical analyses. This approach provided a replicable, standardized, and objective method for evaluating the quality of written academic work across all participants, minimizing potential biases associated with human ratings.

6.8 Mental health

Perceived psychological well-being was assessed using the Ryff Psychological Well-Being Scale (PWBS) (Ryff, 1989). The scale consists of multiple subdimensions (e.g., autonomy, environmental mastery, personal growth, purpose in life), with responses given on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores indicate greater psychological well-being. The PWBS has been widely validated and used across cross-cultural educational contexts (Ryff and Keyes, 1995). Li (2014) tested a shorter version in the Chinese language, and it was used in the present study. In the current sample, internal consistency was acceptable ($\alpha = 0.78$).

6.9 Academic engagement

Academic engagement was measured using the Utrecht Work Engagement Scale—Student Version (UWES-S) (Schaufeli et al.,

2002). This 17-item scale captures three core dimensions of engagement—vigor, dedication, and absorption. Participants rated each item on a 7-point scale from 0 (never) to 6 (always). Total scores were calculated by averaging across all items, with higher scores reflecting greater engagement. The UWES-S has demonstrated strong internal consistency and cross-cultural validity (Schaufeli and Bakker, 2004). The Chinese version developed by Fang et al. (2008) was used. In the present study, the scale demonstrated good internal consistency ($\alpha = 0.89$).

6.10 Covariate: prior academic performance

All main analyses included participants' prior academic performance in a literature-related subject as a covariate. Academic records provided a numerical score on a 100-point scale, reflecting performance in the most recent literature course completed before the intervention. This variable was used to control for potential baseline differences in academic ability related to writing, critical reading, and content familiarity. The scores ranged from 18 to 78, with a mean of 44.37 ($SD = 13.46$), indicating moderate variability across the sample. Controlling for this variable allowed for a more accurate estimation of the intervention effects on the outcome measures.

7 Results

7.1 Descriptive statistics and correlations

Before conducting the main analyses, descriptive statistics and bivariate correlations were calculated for all continuous variables: prior academic performance, text quality, perceived mental health, and academic engagement. Table 1 presents the means and standard deviations for each variable. Pearson's correlation coefficients are presented in Table 2. All correlations were statistically significant at the $p < 0.01$ level. Academic performance was positively correlated with text quality ($r = 0.258$, $p < 0.001$), mental health ($r = 0.329$, $p < 0.001$), and engagement ($r = 0.272$, $p < 0.001$). Text quality also showed moderate positive correlations with mental health ($r = 0.406$, $p < 0.001$) and engagement ($r = 0.280$, $p < 0.001$). The strongest association was observed between mental health and academic engagement ($r = 0.568$, $p < 0.001$), suggesting a meaningful link between students' psychological wellbeing and their engagement with academic tasks.

A series of Univariate Analyses of Covariance (ANCOVA) were conducted to examine the effects of intervention condition on three outcome variables: academic writing quality, perceived mental health, and academic engagement. The independent variable was the type of LLM integration (guided use, unguided use, and control), and prior academic performance was included as a covariate in all models.

TABLE 1 Descriptive statistics and pearson correlations between study variables (*N* = 288).

Variable	<i>M</i>	<i>SD</i>	1	2	3	4
1. Academic performance	44.37	13.46	—			
2. Text quality	123.74	22.86	.258**	—		
3. Mental health	3.82	0.82	.329**	.406**	—	
4. Academic engagement	13.64	2.86	.272**	.280**	.568**	—

N = 288 refers to the total number of participants who completed all measures and were included in the descriptive statistics and Pearson correlations. However, only *N* = 246 participants who passed the manipulation check were retained for the main ANCOVA analyses. *M* = mean; *SD* = standard deviation. ** *p* < 0.01.

7.2 Academic writing quality

The ANCOVA revealed a significant effect of intervention condition on writing quality, $F_{(2,253)} = 789.53$, $p < 0.001$, with a very large effect size ($R^2 \text{ adj} = 0.863$). This indicates that the intervention condition explained approximately 86% of the variance in writing performance, reflecting a strong and practically meaningful impact of structured LLM integration.

Estimated marginal means showed that students in the guided LLM use condition produced significantly higher quality texts ($M = 151.35$, $SE = 0.775$) than those in the unguided use ($M = 129.61$, $SE = 0.770$) and control groups ($M = 109.61$, $SE = 0.700$), as Table 2 shows.

Effect sizes computed with pooled standard deviations showed a very large difference between the guided LLM use and control groups (Cohen's $d = 5.16$), a large difference between the guided and unguided groups ($d = 2.53$), and a large difference between the unguided and control groups ($d = 3.83$). These values highlight the strong impact of guided LLM use on writing performance, and confirm that even unguided use resulted in substantially better outcomes compared to no use. All pairwise comparisons were statistically significant after Bonferroni correction ($p < 0.001$), as Table 3 shows.

7.3 Perceived mental health

The analysis also showed a significant effect of intervention condition on perceived mental health, $F_{(2,253)} = 5.78$, $p = 0.004$, with a small to moderate effect size ($R^2 \text{ adj} = 0.097$). This suggests that nearly 10% of the variability in self-reported mental wellbeing was attributable to the different LLM conditions, indicating a modest yet meaningful contribution of guided use to students' perceived psychological health. Students in the guided use condition reported significantly higher mental health scores ($M = 4.15$, $SE = 0.076$) than the control group ($M = 3.81$, $SE = 0.069$, $p = 0.003$), as Table 4 shows.

As Table 5 shows, the difference between the guided and unguided groups ($M = 3.91$, $SE = 0.076$) approached statistical significance ($p = 0.070$), whereas no significant difference was observed between the unguided and control conditions ($p = 0.984$). Effect size estimates indicated a moderate difference between the guided LLM use and control conditions (Cohen's $d = 0.53$), a small to moderate effect between guided and unguided use ($d = 0.31$), and a negligible effect between unguided use and control ($d = 0.14$).

TABLE 2 Estimated marginal means for academic writing quality (controlling for academic performance).

Intervention	Mean	SE	95% CI lower	95% CI upper
Control	109.61	0.70	108.23	110.99
Unguided LLM use	129.61	0.77	128.10	131.13
Guided LLM use	151.35	0.78	149.82	152.87

Note. Means are estimated marginal means adjusted for the covariate (academic performance).

These findings suggest that only structured use of the LLM led to meaningful psychological benefits.

7.4 Academic engagement

Lastly, the ANCOVA for academic engagement indicated a significant effect of intervention condition, $F_{(2,253)} = 6.70$, $p = 0.001$, with a modest effect size ($R^2 \text{ adj} = 0.101$). This means that around 10% of the variance in engagement was explained by the intervention condition, pointing to a small but practically relevant effect of structured LLM integration on students' involvement in academic activities. Students in the guided LLM use group reported the highest engagement scores ($M = 14.67$, $SE = 0.304$), significantly higher than the control group ($M = 13.17$, $SE = 0.275$, $p < 0.001$), as Table 6 shows.

Although the unguided use group ($M = 13.74$, $SE = 0.302$) scored higher than the control group, this difference was not statistically significant ($p = 0.485$), and the difference between the guided and unguided groups was marginal ($p = 0.093$), as Table 7 shows. For engagement, the contrast between guided use and control yielded a moderate effect size (Cohen's $d = 0.55$), while the effect between guided and unguided use was small to moderate ($d = 0.32$), and the difference between unguided use and control was small ($d = 0.21$). These results indicate that guided integration produced a noticeable improvement in students' involvement, whereas unguided use led to minimal gains.

These results suggest that the guided integration of LLMs can significantly enhance students' academic writing and engagement, and may also support improvements in perceived mental health, compared to both unguided use and no use of LLMs.

Hypothesis 4: Intermediate Outcomes in the Unguided LLM Use Condition

TABLE 3 Pairwise comparisons—academic writing quality.

Comparison	Mean difference (I–J)	SE	p	95% CI Lower	95% CI upper
Guided LLM use—control	41.73	1.05	< 0.001	39.20	44.27
Guided LLM use—unguided use	21.73	1.09	< 0.001	19.11	24.36
Unguided LLM use—control	20.00	1.05	< 0.001	17.48	22.52

SE = standard error; CI = confidence interval. Pairwise comparisons were adjusted using the Bonferroni correction.

TABLE 4 Estimated marginal means for perceived mental health (controlling for academic performance).

Intervention	Mean	SE	95% CI lower	95% CI upper
Control	3.81	0.069	3.67	3.94
Unguided LLM use	3.91	0.076	3.76	4.06
Guided LLM use	4.15	0.076	4.00	4.30

Means are estimated marginal means adjusted for the covariate (academic performance).

Hypothesis 4 proposed that students in the unguided LLM use condition would demonstrate intermediate levels of academic writing quality, perceived mental health, and academic engagement, higher than those in the control group but lower than those in the guided use group.

The results provided partial support for this hypothesis. In terms of academic writing quality, the unguided group ($M = 129.61$, $SE = 0.77$) scored significantly higher than the control group ($M = 109.61$, $SE = 0.70$, $p < 0.001$), but significantly lower than the guided group ($M = 151.35$, $SE = 0.78$, $p < 0.001$). These findings confirm the predicted ordinal pattern in this domain.

However, for perceived mental health, the unguided group ($M = 3.91$, $SE = 0.076$) did not differ significantly from the control group ($M = 3.81$, $SE = 0.069$, $p = 0.984$), although it trended lower than the guided group ($M = 4.15$, $SE = 0.076$), with this comparison approaching statistical significance ($p = 0.070$).

Similarly, regarding academic engagement, the unguided group ($M = 13.74$, $SE = 0.30$) did not significantly differ from the control group ($M = 13.17$, $SE = 0.28$, $p = 0.485$). The difference between the unguided and guided conditions ($M = 14.67$, $SE = 0.30$) was marginal ($p = 0.093$).

These results indicate that while the unguided LLM condition yielded intermediate outcomes for academic writing quality consistent with Hypothesis 4, the same pattern was not statistically supported in perceived mental health and academic engagement.

8 Discussion

8.1 H1: Writing quality enhancement through guided LLM use

The results strongly support Hypothesis 1, demonstrating that students who received structured guidance using LLMs achieved significantly higher academic writing quality than those in both the unguided and control groups. The magnitude of the effect was exceptionally large, underscoring the substantial

educational potential of guided LLM integration. This finding aligns with the growing body of evidence suggesting that effective integration of LLMs in academic contexts improves the quality of student output and encourages critical engagement with both the writing process and the technology itself (Cash et al., 2025). The superiority of the guided condition can be interpreted through several converging mechanisms identified in recent research. First, structured frameworks like the Writing Path, which utilize explicit outlines, have been shown to significantly improve text generation quality by aligning outputs with the user’s intentions and task-specific goals (Lee et al., 2024). This alignment is particularly important in academic settings, where coherence, argument structure, and adherence to conventions are critical.

Moreover, the results reflect broader findings in human-AI collaborative writing research. Studies on tasks such as headline generation show that users achieve better outcomes when they can guide and selectively refine LLM outputs. This process enhances quality without compromising user agency or perceived authorship (Ding et al., 2023). This suggests that guided LLM use in educational settings may strike a productive balance between automation and student ownership. At a cognitive level, guided use of LLMs appears to support key phases in the writing process, particularly translation and revision. Chakrabarty et al. (2024) found that professional writers benefited most from LLM support during these stages. This insight resonates with our results and further substantiates the utility of guided approaches in educational contexts.

Finally, it is worth noting that the enhanced writing performance observed may not stem solely from the tool’s linguistic capabilities, but also from reduced uncertainty and cognitive load due to structured task framing. When students know exactly how to proceed and what is expected of them in using a complex tool like an LLM, their cognitive resources may be more efficiently allocated to higher-order writing concerns, thus improving final output quality.

8.2 H2: Guided LLM use and perceived mental health

The findings provide empirical support for Hypothesis 2, indicating that students in the guided LLM use condition reported significantly higher levels of perceived mental health compared to those in the control group. Although the effect size was modest, the statistical significance of the difference underscores the potential of guided LLM integration as a psychologically beneficial educational tool. Notably, the comparison between the

TABLE 5 Pairwise Comparisons – Mental Health.

Comparison	Mean difference (I–J)	SE	p	95% CI Lower	95% CI Upper
Guided LLM use—control	0.35	0.10	.003	0.10	0.59
Guided LLM use—unguided Use	0.24	0.11	0.070	−0.01	0.50
Unguided LLM use—control	0.10	0.10	0.984	−0.35	0.15

SE, standard error; CI, confidence interval. Pairwise comparisons were adjusted using the Bonferroni correction.

TABLE 6 Estimated marginal means for academic engagement (controlling for academic performance).

Intervention	Mean	SE	95% CI lower	95% CI upper
Control	13.17	0.28	12.62	13.71
Unguided LLM use	13.74	0.30	13.15	14.34
Guided LLM use	14.67	0.30	14.07	15.27

Means are estimated marginal means adjusted for the covariate (academic performance).

guided and unguided groups approached significance, suggesting that guidance in LLM use may play a decisive role in how such tools influence users’ well-being. These results are consistent with growing evidence that conversational AI systems can enhance users’ subjective mental health experiences when implemented with structured guidance. One contributing factor may be the enhanced user experience associated with anthropomorphically designed systems. For instance, Wu et al. (2024) showed that agents like Sunnie increased users’ perceptions of usability and engagement. Such design strategies may foster a more human-like, empathetic interaction, which resonates with students in high-stress academic contexts.

Beyond surface-level interaction quality, systems like VITA have demonstrated that adaptive, behavior-sensitive guidance improves not just perception but also outcomes in mental well-being (Spitale et al., 2025). These systems personalize responses to individual user profiles and evolving needs, features that align well with the nature of guided LLM use in this study. When students receive structured prompts, reflective exercises, or scaffolded interactions from LLMs, the result is improved engagement and potentially heightened psychological support.

The results also echo the broader literature on LLM-based agents such as Replika, which offer on-demand, judgment-free interactions. Ma et al. (2023) highlighted how such interactions help individuals engage in self-reflection and develop confidence. In the present context, the structured engagement with LLMs may serve a similar purpose, providing students with an emotionally neutral space to articulate their thoughts and manage academic stress more effectively. Furthermore, Kumar et al. (2024) noted that AI-guided systems can integrate multiple data sources to detect subtle shifts in mental states and deliver personalized micro-interventions. Although this study did not leverage multimodal inputs, the positive outcome observed in the guided condition suggests that even text-based interventions, when strategically framed, can produce a meaningful uplift in well-being.

At the same time, the non-significant difference between the unguided and control groups raises important questions about the boundary conditions under which LLMs can support mental health. One plausible explanation is that unguided access may generate uncertainty, confusion, or even decision fatigue when students are left to navigate the system without structure. Prior research suggests that the absence of guidance can lead to overwhelming interactions or passive use of the tool, which may fail to produce affective benefits (Zhang et al., 2024; Zhu, 2024). It is possible that psychological support through AI requires not only access but also a sense of clarity, safety, and intentionality—conditions more likely to be fostered in structured interventions. In sum, the significant improvement in perceived mental health among students in the guided condition supports the hypothesis that structured, intentional interaction with LLMs can enhance psychological experiences in academic settings. These findings reinforce the view that LLMs—when deployed thoughtfully—can act as supportive companions in educational environments (Youn and Jin, 2021), particularly when combined with design principles and adaptive features that foster trust, personalization, and emotional safety (Liu et al., 2023). However, the lack of improvement in the unguided condition highlights the importance of pedagogical framing as a necessary condition for translating technological affordances into emotional gains.

8.3 H3: Guided LLM use and academic engagement

The results support Hypothesis 3, indicating that students in the guided LLM use condition exhibited significantly greater academic engagement than those in the control group. The ANCOVA revealed a statistically significant effect of condition on engagement scores, with a modest effect size. Students in the guided condition reported the highest levels of engagement, reinforcing the view that structured interaction with LLMs can foster a more involved and focused learning experience. These findings align with a growing body of research highlighting the importance of guidance in shaping students’ cognitive and emotional engagement in AI-supported learning environments. In particular, structured guidance during LLM use has been shown to reduce off-task behavior, such as random or superficial queries and the indiscriminate use of AI for answer retrieval (Kumar et al., 2024, 2023). By promoting intentional and reflective engagement, guided LLM interventions encourage students to assume more active roles in their learning processes.

TABLE 7 Pairwise comparisons—academic engagement.

Comparison	Mean difference (I–J)	SE	<i>p</i>	95% CI lower	95% CI upper
Guided LLM use—control	1.51	0.41	< 0.001	0.51	2.50
Guided LLM use—unguided use	0.93	0.43	0.093	−0.10	1.96
Unguided LLM use—control	0.58	0.41	0.485	−0.41	1.57

SE, standard error; CI, confidence interval. Pairwise comparisons were adjusted using the Bonferroni correction.

Moreover, research on immersive and AI-integrated educational formats—such as Alternate Reality Games (ARGs) augmented with LLM guidance—has demonstrated that such designs can enhance behavioral engagement, emotional connection, and control beliefs (Cheng et al., 2022; Neary and Schueller, 2018). These findings suggest that when LLMs are embedded in pedagogically grounded frameworks, they can catalyze sustained academic motivation and participation. At the cognitive level, integrating LLMs into virtual teaching assistant roles, such as the Jill Watson system, further supports the idea that guided AI interactions can promote higher-order thinking and intellectual curiosity (Maiti and Goel, 2024). Students in such systems are not merely passive recipients of information but are actively encouraged to formulate and refine complex inquiries, fostering deeper engagement with content.

In contrast, using unguided LLMs favors quick information retrieval over sustained learning. Although such interactions may yield short-term performance gains, they do not appear to generate the same level of student investment or trust in the learning process (Kumar et al., 2025). This may help explain why the guided condition in the present study outperformed both the unguided and control groups regarding engagement. Indeed, the absence of clear instructional framing in the unguided condition may have led to uncertainty about how to use the tool productively, diluting its potential benefits for emotional or behavioral engagement. When students are unsure whether they are using a tool “correctly,” this ambiguity can undermine their sense of efficacy and reduce motivation to persist. Thus, although the results confirmed that guided LLM use fosters greater academic engagement, they also suggest that without supportive structure, LLMs may not reliably elicit active academic involvement. By combining technological capabilities with pedagogical intentionality, these systems offer an interactive and supportive environment that encourages students to actively participate, reflect, and persist in their academic work.

8.4 H4: Intermediate outcomes of unguided LLM use

Hypothesis 4 posited that students in the unguided LLM use condition would exhibit intermediate levels of academic writing quality, perceived mental health, and engagement, greater than those in the control group but lower than those in the guided use condition. The results partially supported this hypothesis: while this expected pattern was observed and statistically confirmed in

academic writing quality, it was not replicated in perceived mental health or academic engagement.

The writing results suggest that access to LLMs can meaningfully enhance students’ output even without structured guidance. The unguided group significantly outperformed the control group, indicating that basic interaction with the tool—through prompts, content generation, or surface-level feedback—was sufficient to raise writing quality. This aligns with prior findings that when used independently, LLMs can offer valuable assistance in planning, drafting, and refining text (Jungheer, 2023; Meyer et al., 2024). Nevertheless, the superior performance of the guided group supports the idea that structured scaffolding, metacognitive prompts, and explicit instruction in tool usage amplify these benefits (Cash et al., 2025; Salimi and Hajinia, 2025).

In contrast, the data did not support the predicted intermediate pattern for perceived mental health. Although the unguided group reported slightly higher scores than the control group, this difference was not statistically significant. One possible explanation is that unguided use, while offering on-demand support, lacks the emotional structure and stress regulation strategies typically embedded in guided implementations. Without clear boundaries or reassurance about responsible usage, students may experience friction, uncertainty, or even anxiety about whether they are “using the tool correctly,” which may counteract any potential gains in psychological well-being (Zhang et al., 2024). In this context, guidance may not only clarify functionality but also serve a regulatory role—normalizing AI integration, reducing confusion, and promoting a sense of support and competence (Zhu, 2024).

A similar pattern emerged with academic engagement. Although the unguided group showed numerically higher engagement than the control group, this difference was not statistically significant. This suggests that, while unguided LLM access may spark curiosity and enable autonomous exploration, it does not consistently produce sustained or deep engagement. One likely reason is that students without instructional scaffolding may remain uncertain about how to engage productively with the tool, leading to hesitant or fragmented interaction. Prior research indicates that without pedagogical framing, students may use LLMs for surface-level information retrieval or task avoidance, limiting the depth of their involvement (Chen and Leitch, 2024). By contrast, guided use has been associated with stronger emotional and cognitive engagement, as students are trained to leverage LLMs in a reflective and goal-oriented manner (Beurer-Kellner et al., 2024; Uchendu et al., 2023).

Altogether, the findings highlight the nuanced role of guidance in realizing the potential of LLMs. While unguided use may yield

modest benefits in writing quality, its impact on mental health and engagement appears more limited. The lack of significant differences between the unguided and control groups in two of the three outcome domains suggests that students may underutilize these tools or even encounter friction in their use without structured scaffolding. These results point to the importance of not only providing access to AI tools but also offering appropriate pedagogical frameworks to ensure their effective and psychologically supportive implementation.

8.5 Limitations

Despite this experimental design's strengths, several limitations should be acknowledged. First, although random assignment was used to allocate participants across conditions, the study relied on self-report manipulation checks to confirm adherence to the assigned use of LLMs. While this approach ensured consistency between reported and intended use, it may introduce response bias or fail to capture subtle variations in how students interpreted and used the tool within each condition (Roshanaei, 2024). Although technically feasible alternatives such as behavioral logging (e.g., tracking LLM interactions) might offer more objective verification, such data were not collected due to ethical constraints and institutional limitations in access control. Future research could explore ways to integrate such measures in a transparent and privacy-respecting manner.

Second, the study was conducted within a single institutional context and with a relatively homogeneous sample—undergraduate students enrolled in humanities and arts programs at an international university in China. This limits the generalizability of the findings to other educational settings, disciplines, and cultural contexts (Salimi and Hajinia, 2025). In particular, students in STEM fields might interact with LLMs differently not only due to varying levels of digital literacy, but also because of the nature of the writing tasks they face, the disciplinary conventions they follow, and the specific modes of information retrieval their fields require. These differences may influence how beneficial, usable, or trustworthy LLM tools appear in practice.

Third, the primary measure of writing quality relied on the automated Performance Score generated by the Grammarly for Education platform. While this tool offers objectivity and replicability, it prioritizes surface-level features such as grammar, clarity, and lexical variety. As a result, it may not fully capture deeper dimensions of academic writing—such as argumentation structure, critical analysis, originality, synthesis of sources, or adherence to disciplinary conventions—which are essential in evaluating high-level academic work. Human-rated assessments or rubric-based evaluations could complement automated scoring in future studies to provide a more holistic picture of writing quality (Salimi and Hajinia, 2025).

Fourth, although the study included a validated measure of prior academic performance as a covariate, this measure was limited to students' most recent literature course. While this represents a meaningful control, the category of "literature-related subject" remains broad and may encompass varying levels of complexity and assessment standards. Moreover, other potentially

relevant factors—such as writing experience in other languages, previous exposure to AI tools, or individual motivation—were not controlled and could have influenced the outcomes (Xu et al., 2025). Fifth, the study used perceived mental health and academic engagement as outcome variables measured through self-report scales. While these instruments are widely validated, self-reported data are subject to social desirability and may not accurately reflect behavioral engagement or psychological functioning (Meyer and Elswiler, 2025). Future research could enhance robustness by incorporating behavioral (e.g., time-on-task) or physiological (e.g., stress monitoring) indicators to triangulate self-perceptions with observable evidence (Youn and Jin, 2021).

To sum up, the findings should be interpreted with an awareness of existing limitations. For example, while guided LLM use markedly improves performance on structured academic tasks, its efficacy may vary across genres or in more creative domains. Gómez-Rodríguez and Williams (2023) noted that human writers still maintain an advantage in areas such as humor and originality, which are difficult for LLMs to replicate reliably. Therefore, while our results highlight the transformative potential of guided LLM use, they also reinforce the need for human oversight and creative judgment in academic writing.

8.6 Practical implications for teaching: structured integration of LLMs in academic instruction

Integrating LLMs into academic writing instruction presents promising opportunities and pressing challenges. The present findings reinforce the importance of structured, guided use of LLMs, particularly in enhancing students' writing quality, academic engagement, and, to a certain extent, their perceived mental well-being. These results carry several practical implications for educators, instructional designers, and policymakers in higher education. Importantly, the implications presented here are directly informed by the limitations discussed above, and their placement after the limitations section reflects a deliberate decision to ensure that recommendations are realistic, context-aware, and attuned to the boundaries of the current design.

8.7 Designing guided LLM integration

The study underscores the pedagogical value of structured engagement with LLMs. Educators should prioritize guided frameworks when introducing LLMs into learning environments (Alsobeh and Woodward, 2023). This includes providing students with clear instructions on using these tools effectively, offering structured prompts, and integrating LLM interactions into existing learning goals (Chiang and Lee, 2023). As demonstrated by approaches such as the Writing Path framework (Lee et al., 2024), guided strategies help align LLM-generated content with academic standards and user intentions, improving writing quality (Gómez-Rodríguez and Williams, 2023). Guidance also plays a crucial role in shaping student behavior during AI interaction. Research shows that structured guidance can reduce off-task or random queries

and foster a more focused, problem-solving approach to writing (Kulaksiz, 2024). When embedded in pedagogy, these strategies enhance performance and increase students' trust and sense of ownership in the learning process (Bekker, 2024).

8.8 The role of educators in mediating LLM use

LLM integration necessitates rediscovering the educator's role—from transmitter of knowledge to AI literacy and ethical engagement facilitator (Lazebnik and Rosenfeld, 2024). Teachers should take an active role in helping students understand the limitations of LLMs, differentiate between responsible use and misuse, and navigate the ethical considerations associated with AI-generated content (Lee, 2023). Training students to critically evaluate and revise LLM outputs contributes to deeper learning and helps prevent overreliance (Liao et al., 2023). Instructors can also promote transparency by encouraging students to document their use of LLMs in the writing process, thus reinforcing principles of academic integrity and accountability (Mahmoud and Sørensen, 2024). This approach fosters a culture of AI-augmented authorship, where students learn to integrate feedback rather than delegate writing tasks to an automated agent. This pedagogical vision also aligns with global policy agendas—such as the UNESCO Futures of Education framework and the Sustainable Development Goal 4 (SDG4)—by promoting inclusive, equitable, and future-ready higher education that incorporates responsible AI use.

8.9 Risks of unguided use and the need for policy

While the study found that even unguided use of LLMs can yield some benefits, particularly in writing quality, such benefits are significantly more limited without instructional scaffolding. Unguided use has several risks, including superficial engagement, skill stagnation, and academic integrity concerns. Without explicit instruction, students may bypass the cognitive and metacognitive processes essential to writing, relying instead on the fluency of LLMs to complete tasks (Lopes et al., 2024). Moreover, the indistinguishability of AI-generated text from human-authored work poses significant challenges for evaluation (De Villiers et al., 2024; Reinhart et al., 2024). This complicates the role of assessment and highlights the urgent need for institutional policies that address transparency, disclosure practices, and acceptable uses of generative AI in coursework. As noted in the limitations, behavioral metrics and clearer definitions of disciplinary norms could inform these policies, especially when automated scoring tools like Grammarly are involved.

8.10 Fostering independent skill development

Finally, educators must balance leveraging the benefits of LLMs and fostering independent writing skills (Patac and Patac Jr., 2025). While guided use can accelerate learning and reduce

barriers, overdependence on AI tools may inhibit students' ability to think critically and write autonomously (Ouwehand et al., 2025). Integrating LLMs should not replace traditional instruction but complement it through strategy-based interventions, peer review, and scaffolded writing tasks (Li et al., 2023; Lin et al., 2024). Future research might also explore how these strategies can be adapted to STEM disciplines or interdisciplinary programs, as differences in writing genres and task complexity may shape how students engage with LLMs. In line with SD4/s commitment to inclusive and contextually sensitive education, such differentiated approaches are essential to ensuring that AI-enhanced instruction serves diverse learners effectively.

9 Conclusion

This study provides empirical evidence for the differential effects of guided and unguided integration of LLMs on students' academic writing quality, perceived mental health, and academic engagement in higher education. The findings demonstrate that guided LLM use consistently outperforms both unguided and no use, particularly in improving writing outcomes and fostering student engagement. These effects are most pronounced when LLMs are embedded within structured instructional frameworks that promote critical interaction, strategic thinking, and responsible use.

Notably, unguided LLM use yielded only partial benefits. While it enhanced academic writing quality relative to the control group, it did not significantly improve perceived mental health or engagement. These results suggest that unstructured exposure to generative AI may not be sufficient to produce holistic educational gains. Instead, pedagogical scaffolding and active instructor involvement appear essential to unlock the full potential of LLMs in supporting learning, well-being, and student agency. The study contributes to the growing literature on human–AI collaboration in education by underscoring the importance of designing intentional, ethically informed, and learner-centered approaches to AI integration. As educational institutions increasingly adopt LLM-based tools, the distinction between guided and unguided use becomes pedagogically relevant—because of its impact on learning quality, engagement, and wellbeing—and ethically imperative, as it directly affects student autonomy, academic integrity, and equitable access to meaningful AI-supported education.

Looking ahead, future research should explore how guidance strategies can be tailored to different learning profiles, disciplines, and institutional cultures. Longitudinal and mixed-method designs may further illuminate the evolving relationship between students and AI, providing insights into how LLMs shape academic development. Ultimately, the challenge lies in providing access to powerful technologies and designing meaningful and equitable frameworks for their use—frameworks that preserve learning integrity while embracing innovation. The findings of this study call for a thoughtful and pedagogically grounded integration of LLMs in higher education. When embedded in structured learning environments, guided use can enhance academic outcomes, support student wellbeing, and promote ethical use of AI. To realize these benefits, educators must assume an active role in designing, modeling, and monitoring AI engagement, ensuring that LLMs serve as tools for empowerment rather than substitution. Achieving

this vision also depends on ongoing professional development for educators, who must be equipped not only with technical competencies but also with pedagogical strategies to guide students in critically and ethically navigating AI-supported academic tasks.

10 Declarations

Ethical approval and Informed Consent for participation: This study was reviewed and approved by the Institutional Review Board (IRB) of Xi'an International University following the Declaration of Helsinki. Ethical approval was granted on [January 15th, 2024], under the reference number [approval ID, IRB/24/072-HUMARTS]. Before participation, all students received an information sheet outlining the purpose of the study, the nature of the tasks, the voluntary nature of their participation, and their right to withdraw at any point without penalty. Written informed consent was obtained from all participants. The consent form emphasized that participation was anonymous, data would be kept confidential, and results would be used solely for academic research. Participants were also informed that using the LLM was part of an experimental educational intervention and that their course grades would not be affected by their responses or participation.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/htejmf?view_only=54624dbd9f11467ea26242bae037e713lt;bgt;.

Ethics statement

The studies involving humans were approved by Institutional Review Board (IRB) of Xi'an International University following the Declaration of Helsinki. Ethical approval was granted on [January 15th, 2024], under the reference number [approval ID, IRB/24/072-HUMARTS]. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

References

- Adams, R. (2024) *The New Empire of AI: The Future of Global Inequality*. London: Polity Press.
- Alsobeh, A., and Woodward, B. (2023). "AI as a Partner in Learning: a Novel Student-in-the-Loop Framework for Enhanced Student Engagement and Outcomes in Higher Education" *Proceedings of the 24th Annual Conference on Information Technology Education*, Marietta, GA, USA. doi: 10.1145/3585059.3611405
- Ayeni, O. O., Al Hamad, N. M., Chisom, O. N., Osawaru, B., and Adewusi, O. E. (2024). AI in education: a review of personalized learning and educational technology. *GSC Adv. Res. Rev.* 18, 261–271. doi: 10.30574/gscarr.2024.18.2.0062
- Bekker, M. (2024). Large language models and academic writing: five tiers of engagement. *South Afr. J. Sci.* 120, 1–5. doi: 10.17159/sajs.2024/17147
- Beurer-Kellner, L., Fischer, M., and Vechev, M. (2024). "Guiding LLMs the right way: fast, non-invasive constrained generation," in *Proceedings of the 41st International Conference on Machine Learning (Vienna)*, 3658–3673.
- Cash, T. N., Oppenheimer, D. M., and Connell Pensky, A. E. (2025). *You've Got AI Friend in Me: LLMs as Collaborative Learning Partners*. doi: 10.31219/osf.io/8q67u_v3
- Chakrabarty, T., Padmakumar, V., Brahman, F., and Muresan, S. (2024). Creativity support in the age of large language models: an empirical study involving professional writers *Proceedings of the 16th Conference on Creativity and Cognition, Chicago: USA*. doi: 10.1145/3635636.3656201
- Chang, C.-H. (2024). *A Study on the Impact of Two AI-Powered Writing Assistants on EFL Learners' Writing* (ProQuest Dissertations and Theses), National Taiwan Normal University, Taiwan.

Author contributions

MZ: Conceptualization, Resources, Visualization, Validation, Formal analysis, Funding acquisition, Project administration, Writing – original draft, Data curation, Supervision, Investigation, Methodology, Writing – review & editing, Software.

Funding

The author declares that no financial support was received for the research and/or publication of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author declares that no Gen AI was used in the creation of this manuscript. Generative AI was used in the preparation of this manuscript for revision of the language and the sentence structure. All the other content (data analyses and theoretical proposals) were written by the authors.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Chen, C., and Leitch, A. (2024). *LLMs as Academic Reading Companions: Extending HCI through Synthetic Personae*. CHI 2024 Workshop: Challenges and Opportunities of LLM-Based Synthetic Personae and Data in HCI, Honolulu, HI.
- Cheng, X., Zhang, X., Cohen, J., and Mou, J. (2022). Human Vs. AI: understanding the impact of anthropomorphism on consumer response to chatbots from the perspective of trust and relationship norms. *Inform. Process. Manag.* 59:102940. doi: 10.1016/j.ipm.2022.102940
- Chiang, C. H., and Lee, H. Y. (2023). "Can large language models be an alternative to human evaluations?" in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics* (Vol. 1: Long Papers), 15607–15631.
- De La Paz, S. (2005). Effects of historical reasoning instruction and writing strategy mastery in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology*, 97(2), 139–156. doi: 10.1037/0022-0663.97.2.139
- De Villiers, C., Dimes, R., and Molinari, M. (2024). How Will AI Text Generation and Processing Impact Sustainability Reporting? *Critical analysis, a conceptual framework and avenues for future research*. *Sustain. Account. Manag. Policy J.* 15, 96–118. doi: 10.1108/SAMPJ-02-2023-0097
- Deci, E. L., and Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11, 227–268. doi: 10.1207/S15327965PLI1104_01
- Ding, Z., Smith-Renner, A., Zhang, W., Tetreault, J., and Jaimes, A. (2023). Harnessing the power of LLMs: evaluating human-AI Text co-creation through the lens of news headline generation. Findings of the Association for Computational Linguistics: EMNLP 2023, *The 2023 Conference on Empirical Methods in Natural Language Processing*. Singapore. doi: 10.18653/v1/2023.findings-emnlp.217
- Fang, L. T., Shi, K., and Zhang, F. (2008). Research on reliability and validity of Utrecht work engagement scale-student. *Chinese J. Clin. Psychol.* 16, 618–620.
- Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G* power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149
- Flesch, R. (1948). A new readability yardstick. *J. Appl. Psychol.* 32, 221–233. doi: 10.1037/h0057532
- Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Rev. Educ. Res.* 74, 59–109. doi: 10.3102/00346543074001059
- Fritz, T., González Cruz, H., Janke, S., and Daumiller, M. (2024). How to best measure academic dishonesty in students: a systematic review of self-report assessment methods and psychometric quality. *Eur. J. Psychol. Assess.* 40, 498–514. doi: 10.1027/1015-5759/a000861
- Gómez-Rodríguez, C., and Williams, P. (2023). "A confederacy of models: a comprehensive evaluation of LLMs on creative writing," in *Findings of the Association for Computational Linguistics: EMNLP 2023*, 14504–14528, Singapore. Association for Computational Linguistics. doi: 10.18653/v1/2023.findings-emnlp.966
- Granieri, A., Franzoi, I. G., and Chung, M. C. (2021). Editorial: psychological distress amongst university students. *Front. Psychol.* 12:647940. doi: 10.3389/fpsyg.2021.647940
- Jungherr, A. (2023). *Using Chatgpt and Other Large Language Model (LLM) Applications for Academic Paper Assignments*. Bamberg: Otto-Friedrich-Universität. doi: 10.31235/osf.io/d84q6
- Kulaksiz, G. C. (2024). *Artificial Intelligence-Based Language Modelling: The Effect of Chatgpt Application on Writing Skills in the Context of Teaching English as a Foreign Language*. Istanbul: Bursa Uludağ Üniversitesi.
- Kumar, A., Prol, D., Alipour, A., and Ragavan, S. S. (2025). To google or To ChatGPT? A comparison of CS2 students' information gathering approaches and outcomes. *arXiv preprint arXiv:2501.11935*.
- Kumar, H., Musabirov, I., Reza, M., Shi, J., Wang, X., Williams, J. J., et al. (2023). Impact of guidance and interaction strategies for LLM use on Learner Performance and perception. *arXiv preprint arXiv:2310.13712*.
- Kumar, H., Reza, M., Mitchell, J., Musabirov, I., Zhang, L., and Liut, M. (2024). Understanding help-seeking behavior of students using LLMs Vs. web search for writing sql queries. *PsyArxiv preprint*. doi: 10.1145/3735091.3737569
- Lazebnik, T., and Rosenfeld, A. (2024). Detecting LLM-assisted writing in scientific communication: are we there yet? *J. Data Inform. Sci.* 9. doi: 10.2478/jdis-2024-0020
- Lee, S.-M. (2023). The effectiveness of machine translation in foreign language education: a systematic review and meta-analysis. *Comput. Assis. Lang. Learn.* 36, 103–125. doi: 10.1080/09588221.2021.1901745
- Lee, Y., Ka, S., Son, B., Kang, P., and Kang, J. (2024). Navigating the path of writing: outline-guided text generation with large language models. *PsyArxiv preprint*. doi: 10.18653/v1/2025.naacl-industry.20
- Li, C., Zhang, M., Mei, Q., Wang, Y., Amba Hombaiah, S., Liang, Y., and Bendersky, M. (2023). *Teach LLMs to Personalize- an Approach Inspired by Writing Education*.
- Li, C., Zhang, Y., Randhawa, A. K., and Madigan, D. J. (2020). Emotional exhaustion and sleep problems in university students: does mental toughness matter? *Personal. Individ. Diff.* 163:110046. doi: 10.1016/j.paid.2020.110046
- Li, R.-H. (2014). Reliability and validity of a shorter chinese version for ryff's psychological well-being scale. *Health Educ. J.* 73, 446–452. doi: 10.1177/0017896913485743
- Liao, W., Liu, Z., Dai, H., Xu, S., Wu, Z., Zhang, Y., Huang, X., Zhu, D., Cai, H., and Li, Q. (2023). Differentiating chatgpt-generated and human-written medical texts: quantitative study. *J. Med. Int. Res. Med. Educ.* 9:e48904. doi: 10.2196/48904
- Lin, H.-C. K., Lu, L.-W., and Lu, R.-S. (2024). Integrating digital technologies and alternate reality games for sustainable education: enhancing cultural heritage awareness and learning engagement. *Sustainability* 16:9451. doi: 10.3390/su16219451
- Lin, X. (2024). Exploring the role of chatgpt as a facilitator for motivating self-directed learning among adult learners. *Adult Learn.* 35, 156–166. doi: 10.1177/10451595231184928
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., and Neubig, G. (2023). Pre-Train, prompt, and predict: a systematic survey of prompting methods in natural language processing. *ACM Comput. Surv.* 55, 1–35. doi: 10.1145/3560815
- Lopes, R. M., Silva, A. F., Rodrigues, A. C. A., and Melo, V. (2024). Chatbots for well-being: exploring the impact of artificial intelligence on mood enhancement and mental health. *Eur. Psychiat.* 67, S550–S551. doi: 10.1192/j.eurpsy.2024.1143
- Lu, Q., Yao, Y., Xiao, L., Yuan, M., Wang, J., and Zhu, X. (2024). Can chatgpt effectively complement teacher assessment of undergraduate students' academic writing? *Assess. Eval. High. Educ.* 49, 616–633. doi: 10.1080/02602938.2024.2301722
- Ma, Z., Mei, Y., and Su, Z. (2023). Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support. *AMIA Sympos.* 2023, 1105–1114.
- Mahmoud, C., and Sørensen, J. (2024). Artificial Intelligence in personalized learning with a focus on current developments and future prospects. *Res. Adv. Educ.* 3, 25–31. doi: 10.56397/RAE.2024.08.04
- Maiti, P., and Goel, A. K. (2024). How do students interact with an LLM-powered virtual teaching assistant in different educational settings? *arXiv preprint arXiv:2407.17429*.
- Meyer, J., Jansen, T., Schiller, R., Liebenow, L. W., Steinbach, M., Horbach, A., and Fleckenstein, J. (2024). Using LLMs to bring evidence-based feedback into the classroom: ai-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Comput. Educ. Art. Intell.* 6:100199. doi: 10.1016/j.caeai.2023.100199
- Meyer, S., and Elsweiler, D. (2025). Llm-based conversational agents for behaviour change support: a randomised controlled trial examining efficacy, safety, and the role of user behaviour. *Int. J. Hum. Comput. Stud.* 200:103514. doi: 10.1016/j.ijhcs.2025.103514
- Mishra, P., and Koehler, M. J. (2006). Technological pedagogical content knowledge: a framework for teacher knowledge. *Teach. Coll. Rec.* 108, 1017–1054. doi: 10.1111/j.1467-9620.2006.00684.x
- Molodyskiy, A., Lewis, T., Kadhum, M., Farrell, S. M., Lemtiri Chelieh, M., Falcão De Almeida, T., Masri, R., Kar, A., Volpe, U., and Moir, F. (2021). Cultural variations in wellbeing, burnout and substance use amongst medical students in twelve countries. *Int. Rev. Psychiat.* 33, 37–42. doi: 10.1080/09540261.2020.1738064
- Neary, M., and Schueller, S. M. (2018). State of the field of mental health apps. *Cogn. Behav. Pract.* 25, 531–537. doi: 10.1016/j.cbpra.2018.01.002
- Ng, D. T. K., Tan, C. W., and Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through chatgpt: a pioneering pilot study. *Br. J. Educ. Technol.* 55, 1328–1353. doi: 10.1111/bjet.13454
- Ouwehand, K., Lespiau, F., Tricot, A., and Paas, F. (2025). Cognitive load theory: emerging trends and innovations. *Educ. Sci.* 15:458. doi: 10.3390/educsci15040458
- Park, H., and Ahn, D. (2024). The promise and peril of chatgpt in higher education: opportunities, challenges, and design implications. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* 271, 1–21. doi: 10.1145/3613904.3642785
- Patac, L. P., and Patac Jr, A. V. (2025). Using chatgpt for academic support: managing cognitive load and enhancing learning efficiency—a phenomenological approach. *Soc. Sci. Human. Open* 11:101301. doi: 10.1016/j.ssaho.2025.101301
- Reinhart, A., Brown, D., Markey, B., Loudenbach, M., Pantusen, K., Yurko, R., and Weinberg, G. (2024). *Do LLMs Write Like Humans? Variation in Grammatical and Rhetorical Styles*. doi: 10.1073/pnas.2422455122
- Robbleto, E., Habashi, A., Kaplan, M.-A. B., Riley, R. L., Zhang, C., Bianchi, L., and Shehadeh, L. A. (2024). Medical Students' Perceptions of an Artificial Intelligence (AI) Assisted Diagnosing Program. *Med. Teach.* 46, 1180–1186. doi: 10.1080/0142159X.2024.2305369
- Roshanaei, M. (2024). Towards best practices for mitigating artificial intelligence implicit bias in shaping diversity, inclusion and equity in higher education. *Educ. Inform. Technol.* 29, 18959–18984. doi: 10.1007/s10639-024-12605-2
- Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037/0003-066X.55.1.68

- Ryff, C. D. (1989). Happiness Is everything, or Is It? *Explorations on the meaning of psychological well-being*. *J. Personal. Soc. Psychol.* 57, 1069–1081. doi: 10.1037/0022-3514.57.6.1069
- Ryff, C. D., and Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *J. Personal. Soc. Psychol.* 69, 719–727. doi: 10.1037/0022-3514.69.4.719
- Salimi, E. A., and Hajinia, M. (2025). *Large Language Models and Academic writing in practice: exploring participants' utilization of generative pretrained transformers during an AI-Assisted Course on Writing Research Papers*. *PsyArxiv preprint*. doi: 10.21203/rs.3.rs-5534554/v1
- Schaufeli, W. B., and Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: a multi-sample study. *J. Organ. Behav.* 25, 293–315. doi: 10.1002/job.248
- Schaufeli, W. B., Martinez, I. M., Pinto, A. M., Salanova, M., and Bakker, A. B. (2002). Burnout and engagement in university students: a cross-national study. *J. Cross Cult. Psychol.* 33, 464–481. doi: 10.1177/0022022102033005003
- Sharma, S., Mittal, P., Kumar, M., et al. (2025). The role of Large Language Models in personalized learning: a systematic review of educational impact. *Discov. Sustain.* 6:243. doi: 10.1007/s43621-025-01094-z
- Spitale, M., Axelsson, M., and Gunes, H. (2025). Vita: a multi-modal llm-based system for longitudinal, autonomous, and adaptive robotic mental well-being coaching. *ACM Transact. Hum. Robot Interact.* 14. doi: 10.1145/3712265
- Sweller, J. (1988). cognitive load during problem solving: effects on learning. *Cogn. Sci.* 12, 257–285. doi: 10.1207/s15516709cog1202_4
- Uchendu, A., Lee, J., Shen, H., Le, T., and Lee, D. (2023). Does human collaboration enhance the accuracy of identifying llm-generated deepfake texts? *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, 11, 163–174. doi: 10.1609/hcomp.v11i1.27557
- Wang, L. (2022). Student intrinsic motivation for online creative idea generation: mediating effects of student online learning engagement and moderating effects of teacher emotional support. *Front. Psychol.* 13:954216. doi: 10.3389/fpsyg.2022.954216
- Wu, S., Han, F., Yao, B., Xie, T., Zhao, X., and Wang, D. (2024). Sunnie: an anthropomorphic LLM-based conversational agent for mental well-being activity recommendation. *arXiv preprint arXiv:2405.13803*.
- Xu, X., Qiao, L., Cheng, N., Liu, H., and Zhao, W. (2025). Enhancing self-regulated learning and learning experience in generative ai environments: the critical role of metacognitive support. *Br. J. Educ. Technol. Early View* doi: 10.1111/bjet.13599
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., and Gašević, D. (2024). Practical and ethical challenges of Large Language Models in education: a systematic scoping review. *Br. J. Educ. Technol.* 55, 90–112. doi: 10.1111/bjet.13370
- Youn, S., and Jin, S. V. (2021). In AI we trust?" the effects of parasocial interaction and technopion vs. luddite ideological views on chatbot-based customer relationship management in the emerging "feeling economy". *Comput. Hum. Behav.* 119:106721. doi: 10.1016/j.chb.2021.106721
- Zhang, O. X., Zhou, S., Geng, J., Liu, Y., and Liu, S. X. (2024). Dr. GPT in campus counseling: understanding higher education students' opinions on LLM-assisted mental health services. *arXiv preprint arXiv:2409.17572*.
- Zhu, Y. (2024). The impact of AI-assisted teaching on students' learning and psychology. *J. Educ. Humanit. Soc. Sci.* 38, 111–116. doi: 10.54097/k7a37d11

Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

