BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

The Efficacy of English Academic Writing Improvement of Doctoral Students of Humanities and Social Sciences in China

Yuwei Huang [1] **, *Xiaohui Liang*** [1]**, *Chang Li*** [2*]

[1] *School of Foreign Studies, University of Science & Technology Beijing, Beijing 100083, China*

[2] *Department of Education, The Catholic University of Korea, Bucheon 14662, Republic of Korea*

ABSTRACT

Over the past 30 years, the expansion of doctoral student enrollment in China has led to a substantial increase in the number of doctoral students specializing in the humanities and social sciences. Nevertheless, the majority of these students predominantly publish in Chinese-language journals, which often results in a deficiency in their ability to write English academic writing. Consequently, identifying efficacious strategies to enhance their English academic writing proficiency is imperative. The emergence of AI tools has introduced both challenges and opportunities for students and educators in writing training. This paper aims to investigate the differentiation in writing feedback provided by AI systems versus human instructors for the group's English academic writing. The findings indicate that the group demonstrates a greater engagement with AI-generated feedback compared to that provided by human instructors. Analysis revealed that human instructors delivered feedback characterized by greater Accuracy and Essential Feedback ($p = 0.023 < 0.05$; $p = 0.013 < 0.05$), whereas AI feedback excelled in offering clearer Direction for Improvement and a more Supportive Tone ($p = 0.000 < 0.05$; $p = 0.009 < 0.05$). No statistically significant disparity was identified between AI and human instructors concerning Criterion-Based Comments ($p = 0.323 > 0.05$). Based on these empirical findings, this paper proposes recommendations for integrating AI and human instructor feedback mechanisms, aiming to enhance the English academic writing skills of doctoral students, a substantial demographic within China that exhibits an urgent need for such training.

Keywords: Writing Feedback; AI; Human Instructors; English Academic Writing; Humanities and Social Sciences

*CORRESPONDING AUTHOR:

Chang Li, Department of Education, The Catholic University of Korea, Bucheon 14662, Republic of Korea; Email: lincoln-burrows@qq.com

ARTICLE INFO

Received: 10 June 2025 | Revised: 16 June 2025 | Accepted: 19 June 2025 | Published Online: 15 July 2025

DOI: https://doi.org/10.30564/fls.v7i7.10444

CITATION

Huang, Y., Liang, X., Li, C., 2025. The Efficacy of English Academic Writing Improvement of Doctoral Students of Humanities and Social Sciences in China. Forum for Linguistic Studies. 7(7): 565–578. DOI: https://doi.org/10.30564/fls.v7i7.10444

1. Introduction

Providing students with formative feedback on their writing is a well-established pedagogical practice that effectively facilitates writing development[1]. However, the considerable time and effort required to provide such feedback, especially to multiple students across multiple classes, can be daunting for educators, potentially deterring some from offering the necessary writing instruction[2, 3]. The pedagogical burden on instructors could potentially be mitigated by integrating AI-based writing feedback tools. Such integration may facilitate an expansion in opportunities for both writing practice and targeted instructional intervention. The recent emergence of highly sophisticated AI technologies, exemplified by generative pre-trained transformer (GPT) large language models such as OpenAI's GPT-4 and Google's PaLM 2, heralds a transformative era for Automated Writing Evaluation (AWE)[4]. These technologies, building upon earlier tools like Grammarly and Pigai, leverage natural language processing to identify errors and suggest improvements. The rapid advancement of such technologies, alongside the rise of AI writing feedback, has paved the way for scalable, timely, and personalized feedback[5–8]. Consequently, researchers and educators are now investigating how AI can be best integrated into educational practices to maximize its benefits[9].

To further understand the impact of AI writing feedback on students, some scholars have explored its influence on academic engagement, finding that AI-driven chatbots positively foster behavioral, cognitive, and emotional engagement among Chinese EFL students[6, 10]. Others have conducted comparative studies on the effects of AI writing feedback and human evaluation, revealing that both have their respective advantages, such as the supportive tone of AI or the accuracy of human assessors[1, 4]. However, the majority of these studies have focused on undergraduates or middle school students, with few targeting doctoral students, particularly those in the humanities and social sciences. In China, while doctoral students in science and engineering disciplines often engage in English academic writing early on, many students in the humanities and social sciences tend to publish in Chinese journals, indicating a need to improve their English academic writing skills[11]. As publication in Chinese core journals becomes increasingly competitive, more doctoral students in the humanities and social sciences are seeking publication in international indexes such as the Social Sciences Citation Index (SSCI) and the Arts & Humanities Citation Index (A&HCI), which further highlights their deficiencies in English academic writing[11]. This study seeks to conduct a comparative analysis of academic engagement and the efficacy of AI-generated writing feedback versus human instructor feedback among Chinese doctoral students in the humanities and social sciences. The investigation aims to identify the distinct advantages inherent in each feedback modality, thereby informing pedagogical strategies designed to enhance the English academic writing proficiency of this demographic. Furthermore, the research endeavors to elucidate how human evaluators might adapt their practices in response to the evolving landscape of AI writing feedback.

2. Academic Engagement in Writing

The notion of "academic engagement" encompasses a comprehensive spectrum of student behaviors and attitudes, signifying the depth and intensity of their participation or commitment to scholarly activities[12]. This conceptual framework delineates the degree to which learners are actively and enthusiastically immersed in the learning process, encompassing cognitive, emotional, and behavioral dimensions of their involvement in the academic domain[13]. Empirical investigations conducted to date have substantiated the premise that the students' cognitive, behavioral, and emotional engagement significantly enhances their language acquisition outcomes, including, but not limited to, the proficiency and efficacy of their academic writing[14–16]. Specifically, in the context of second language education, particularly within the realm of academic writing in English, the extension of the concept of academic engagement necessitates a nuanced and multifaceted approach. In the seminal work, Sang and Hiver (2021) have elaborated upon this construct by delineating it into three interrelated and integral components: emotional, cognitive, and behavioral engagement[13]. These dimensions collectively capture the multifaceted nature of student involvement (academic engagement) in the complex process of second language acquisition (SLA), such as academic writing.

2.1. Emotional Engagement in Academic Writing

Emotional engagement constitutes a pivotal dimension within the realm of language acquisition, encapsulating the intricate spectrum of affective responses and motivational dispositions that learners manifest throughout their immersion in the language learning experience. For instance, a supportive tone in writing feedback catalyses fostering a positive emotional climate, which, in turn, nurtures learners' intrinsic motivation and encourages them to engage deeply with the subject matter. Effective feedback, at its essence, is characterised by a combination of affirmatory elements that serve to reinforce positive behaviors and outcomes. It is conveyed through a linguistic style that maintains a supportive tone, and such feedback is structured in a non-directive manner[17], allowing the recipient to retain autonomy in the learning process and to internalize the insights provided. As Hattie and Timperley point out, student self-efficacy is key in responding to feedback, especially when it is supportive or not[18, 19]. Moreover, a study found that underscoring the significance of positive feedback as a catalyst for behavioral transformation is paramount to fostering an environment conducive to enhancing students' learning outcomes[20]. This approach not only reinforces desirable actions but also cultivates a mindset in learners that their efforts are recognized and valued, which in turn stimulates intrinsic motivation and encourages a cycle of continuous academic growth and personal development. Therefore, Emotional engagement can be understood as the affective responses and motivational dispositions that learners exhibit towards the language learning experience, such as a supportive tone in writing feedback for academic writing.

2.2. Cognitive Engagement in Academic Writing

Cognitive engagement delves into the realm of mental and intellectual endeavors that learners undertake in their quest to master the target language, particularly within the domain of academic writing. Central to this cognitive engagement are criteria-based comments and accurate feedback, which serve as essential tools for guiding learners' intellectual development. Criteria-based comments provide a structured framework for evaluating learners' work, allowing them to understand the specific standards and expectations inherent in academic writing. Criteria-based comments entail elucidating the specific benchmarks against which performance is evaluated, thereby providing learners with a clear understanding of expectations and facilitating targeted improvements, which is a key characteristic of effective feedback[21, 22]. By explicitly articulating these criteria in learning such as academic writing, educators provide students with a roadmap for success, enabling them to critically evaluate their work, identify areas for enhancement, and develop strategies to bridge the gap between current performance and desired outcomes, therefore, students can better understand their progress toward successful writing in a genre[17].

Research underscores the importance of explicitly stating criteria for enhancing student writing skills, noting that the effectiveness of writing rubrics is maximized when they cleary convey unambiguous expectations and criterion-based delineations of superior writing[23]. While accurate feedback is crucial, the importance of accuracy highlights that feedback which is imprecise or erroneous can engender significant confusion or even foster disengagement in students during the crucial revision process[24]. It accentuates that for feedback to guide learning and promote improvement effectively, it must be rooted in clear, precise, and reliable information that accurately reflects the strengths and weaknesses of a student's work. Conversely, inaccurate feedback not only fails to provide valuable direction but may also exacerbate misunderstandings, thereby diminishing the learner's motivation to revise and refine their work. In a research study investigating students' perceptions of Automated Writing Evaluation (AWE) within an educational setting, the erroneous feedback generated by AWE necessitated the involvement of human experts to guide students towards more targeted and relevant feedback for their attention[25]. Hence, the accuracy of feedback is a cornerstone in the architecture of effective educational practices, ensuring that the learning journey is both productive and conducive to the academic growth of writing.

2.3. Behavioral Engagement in Academic Writing

Behavioral engagement emerges as a discernible facet of the language learning process, characterized by the observable actions and participation patterns that learners exhibit.

This dimension is vividly exemplified through clear actions aimed at enhancing academic writing skills. When learners receive feedback that not only identifies areas for improvement but also provides Clarity of Directions for Improvement and Prioritization of Essential Features in academic writing, they are more likely to demonstrate active behavioral engagement. To achieve Clarity of Directions for Improvement, empirical research in the field of education underscores the importance of feedback that is characterized by clarity and the use of precise language[18, 22]. Such feedback is instrumental in provoking actionable writing strategies among students. It is evident that when feedback is delivered with a high degree of clarity, it not only communicates the strengths and weaknesses of the student's writing but also provides a clear pathway for improvement. However, feedback that is nebulous or imprecise is less likely to be comprehended and integrated by authors, particularly in instances where the student possesses an intrinsic low level of self-efficacy about writing[26, 27]. This approach to feedback is deemed effective because it directly contributes to the development of the student's writing skills and enhances their ability to self-regulate and refine their work. While Prioritization of Essential Features in academic writing instruction, it is recognized that an excessive volume of feedback can be daunting and counterproductive for students. Consequently, effective feedback is carefully curated to prioritize those essential aspects of writing that are both attainable and reasonable for a student to concentrate on in their subsequent writing endeavors[25]. This strategic approach ensures that the feedback is not only manageable in scope but also targeted towards the most critical areas for improvement, thereby facilitating a focused and constructive learning process. Suppose the feedback focuses excessively on aspects that are not central to the student's current stage of development or the specific assignment at hand. In that case, it can divert their attention, which would genuinely enhance their writing skills. While feedback that is meticulously designed to prioritize the essential features of writing and concentrate on higher-order concerns is inherently more conducive to supporting the learning process. This type of feedback is particularly effective in fostering educational growth, especially when it is rendered manageable, not only concerning the volume of comments provided but also in alignment with the student's current level of writing proficiency[19, 28].

As discussed above, Academic Engagement in academic writing consists of three key elements: (1) Behavioral Engagement (Clear Direction For Improvement; Essential Features); (2) Cognitive Engagement (Criteria-Based Comments; Accuracy), and (3) Emotional Engagement (Supportive Tone). Therefore, The Framework Of Academic Engagement for the research can be constructed as shown in **Figure 1**.

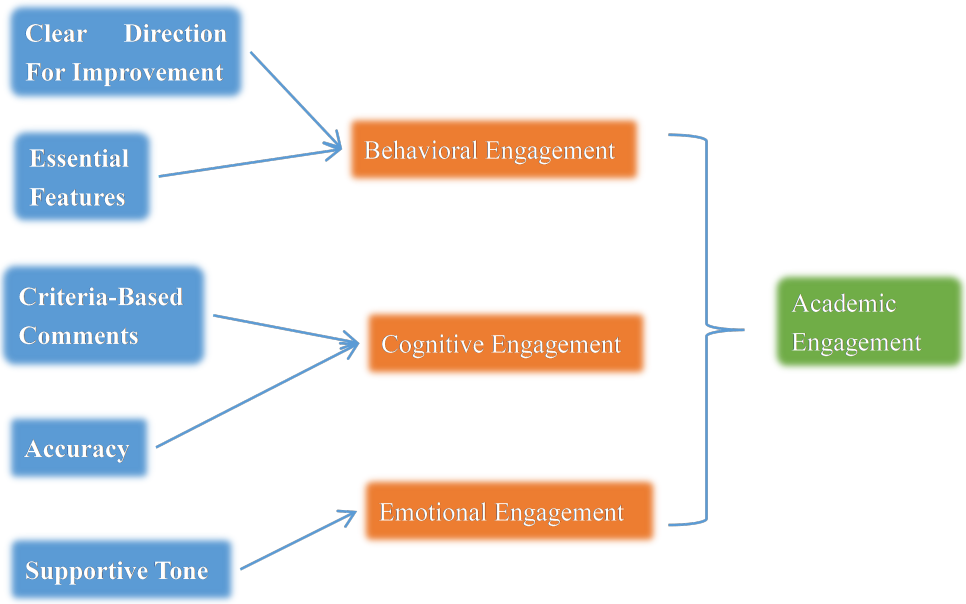

Figure 1. The framework of academic engagement.

Each subsection has been studied from different aspects by various scholars, providing valuable insights into exploration of academic engagement and a comparison of the effectiveness of AI writing feedback with that of human instructors among doctoral students of humanities and social sciences in China. Based on the previous studies, this paper intends to explore the following Research Questions:

RQ1: Does the assistance of AI significantly improve the academic engagement of doctoral students of humanities and social sciences in the publication of English academic journals?

RQ2: The writing feedback provided by AI or writing feedback provided by human evaluators: which helps doctoral students of humanities and social sciences improve more in English article writing from a different perspective?

RQ3: What can we expect with AI in the future enhancement of English article writing for doctoral students of humanities and social sciences?

3. Methodology

3.1. Participants

The 372 newly enrolled doctoral students of humanities and social sciences in Beijing ranged from 22 to 46 years old, who studied in languages/literature/philosophy/management/economies 37.1% of them had the experience of publishing articles in English Academic Journals, among them, 11% of them had published English articles in SSCI/SCI/A&HCI (not as the first author), while 10.9% of them had published English articles in SSCI/SCI/A&HCI (as the first author), the rest of them had published English articles in other journals. Besides the 37.1%, all other participants had only published Chinese articles in Chinese academic journals. More details will be listed in **Table 1.**

Table 1. The Demographics of participants.

Gender	Disciplines	Academic Writing Background
Male: 211	Languages/Literature/Philosophy: 103	Publications in English Journals: 138 (37.1%)
Female: 1611	Management/Economies: 101	Publications in Chinese Journals: 234 (62.9%)
	Others: 168	

3.2. Research Design

Step 1. A pre-test was administered to evaluate the academic writing proficiency of all 372 first-year doctoral students enrolled in humanities and social science programs. Participants were tasked with composing one English academic essay per week (the initial week involved participation in an online preparatory course). A panel of ten writing assessors subsequently assessed these essays. These assessors had undergone prior calibration testing to ensure inter-rater reliability, and their consistent scoring was verified during the assessment process. Based on these evaluations, each student received a proficiency rating, which was then used to stratify the cohort into two groups of equal size and similar writing performance levels: a control group and an experimental group, each comprising 186 students.

Step 2. The control group, consisting of 186 students, received no specific instruction on utilizing AI tools for writing enhancement. While acknowledging the practical impossibility of completely preventing students from accessing AI resources during their time, it is important to note that this group did not receive any systematic training or structured guidance on leveraging AI chatbot feedback for academic writing, unlike the experimental group.

Step 3. Conversely, the experimental group, also comprising 186 students, participated in a teacher-guided intervention. Faculty members provided explicit instruction on the effective use of several prevalent AI chatbot applications widely available in China, such as DeepSeek, Kimi, and Spark. Students were permitted to utilize one or multiple selected AI chatbots concurrently to revise and enhance their academic writing, leveraging the feedback provided by these tools as part of their learning process.

Step 4. Following the completion of six instructional sessions, all participants were required to compose a final English academic article during the concluding week. This corpus of work was then evaluated by the same panel of ten writing assessors, who provided ratings reflecting the

students' overall writing performance. Subsequently, a comparative analysis was conducted to ascertain the differences in performance between the two study groups. Following this analysis, a questionnaire developed by Zhou et al. (2021) was administered to doctoral students of humanities and social sciences disciplines. This instrument was utilized to measure their academic engagement, specifically examining three dimensions previously illustrated in Figure 1: (1) Behavioral Engagement (encompassing "Clear Direction For Improvement" and "Essential Features"), (2) Cognitive Engagement (assessed via "Criteria-Based Comments" and "Accuracy"), and (3) Emotional Engagement (evaluated based on "Supportive Tone"). The investigation also aims to establish a comprehensive framework for researchers and educators, enabling the quantification and deeper understanding of the complex engagement profiles of Chinese doctoral students in humanities and social sciences, particularly in relation to English academic writing. Furthermore, all data gathered from this study will undergo statistical analysis using SPSS version 20.

4. Data Analysis

Initially, a paired-samples t-test was conducted to assess the academic writing improvement of doctoral students in the humanities and social sciences across two groups following a semester-long training program (Group 1 received instruction from human instructors, while AI instructors trained Group 2). Subsequently, an independent samples t-test was employed to ascertain the statistical significance of the differences between the post-test scores of the two groups.

As shown in **Table 2,** there is a statistically significant enhancement in the performance of Group 1 from the pre-test to the post-test phase ($p < 0.05$; Mpre = 66.06, Mpost = 74.02). This indicates that doctoral students of the humanities and social sciences experienced a significant improvement in their academic writing proficiency after receiving feedback from human instructors during the semester.

Table 2. Paired-samples t test of pre- and post-test (group 1).

Group Statistics					
	test	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
VALUE	1	192	66.06	13.844	0.999
	2	192	74.02	10.847	0.783

Test Statistics[a]	
	Values
Mann-Whitney U	12084.000
Wilcoxon W	30612.000
Z	-5.846
Asymp. Sig. (2-tailed)	0.000

[a] Grouping Variable: TEST

As indicated in **Table 3,** the assessment results for Group 2 show a statistically significant improvement in post-test performance compared to pre-test results ($p = 0.000 < 0.05$; Mpre = 65.75, Mpost = 73.23). This suggests that doctoral students in the disciplines of humanities and social sciences have achieved a considerable enhancement in their academic writing skills following the receipt of AI-generated writing feedback over this semester.

The results presented in **Table 4** indicate a statistically insignificant disparity in post-test performance between the two experimental groups ($p = 0.452 > 0.05$; Mean Group 1 = 74.02, Mean Group 2 = 73.23). This finding suggests that doctoral students in the humanities and social sciences within both cohorts experienced a considerable improvement in their academic writing skills following the receipt of feedback, whether delivered by human instructors or through AI writing assistance, over the semester. However, the magnitude of this progress did not differ markedly across the groups, indicating that both feedback modalities were comparably effective in augmenting students' writing proficiency. This convergence of outcomes prompts a critical reflection on the evolving landscape of pedagogy. While human instructors currently hold a dominant position, the study underscores the potential for AI to emerge as a formidable contender in the realm of academic writing

support. The rapid pace of technological advancement may eventually outstrip the capacity of human educators to scale their interventions universally. Furthermore, it is conceivable that AI could introduce novel instructional modules, thereby fundamentally reshaping the ecosystem of writing pedagogy. After the initial phase of data analysis, the second phase employed an independent samples t-test. This statistical procedure was implemented to elucidate doctoral students' perceptions of Academic Engagement—specifically examining the behavioral, cognitive, and affective dimensions—when interacting with human instructors versus AI instructors within the context of English academic writing training.

Table 3. Paired-samples t test of pre- and post-test (group 2).

Group Statistics					
	test	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
VALUE	3	192	65.75	13.580	0.980
	4	192	73.23	10.694	0.772
Test Statistics[a]					
				Values	
Mann-Whitney U				12248.000	
Wilcoxon W				30776.000	
Z				-5.694	
Asymp. Sig. (2-tailed)				0.000	
[a] Grouping Variable: TEST					

Table 4. Independent sample t test of post-test (2 groups).

Group Statistics					
	test	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
VALUE	2	192	74.02	10.847	0.783
	4	192	73.23	10.694	0.772
Test Statistics[a]					
				Values	
Mann-Whitney U				17615.500	
Wilcoxon W				36143.500	
Z				-0.753	
Asymp. Sig. (2-tailed)				0.452	
[a] Grouping Variable: TEST					

As indicated by **Figure 2** and **Table 5**, the analysis of student engagement with AI writing feedback within the context of doctoral education has yielded intriguing insights. The data indicates a marked discrepancy in overall engagement between the two groups, with a statistically significant difference ($p < 0.05$) emerging. Notably, Group 2, which was systematically instructed to utilize AI for writing feedback, exhibited a higher level of engagement ($M = 3.61$) compared to Group 1 ($M = 2.66$). This disparity suggests that the doctoral students in Group 2 have a general affinity for engaging with AI to refine their academic writing skills, an affinity reflected in their elevated mean engagement score, which surpasses the median value of 3, indicating a preference for AI-generated feedback over traditional teacher feedback.

In examining the specific dimensions of engagement, several patterns emerge. Concerning Behavioral Engagement (quantifies the dedication of time and energy that students allocate to academic activities), Group 2 demonstrated a significantly greater commitment of time and energy to academic activities ($p < 0.05$; $M = 3.07$) than Group 1 ($M = 2.06$). This finding underscores the proactive role that AI plays in facilitating student involvement in the writing process.

Similarly, in terms of Emotional Engagement (evaluates the sentiments and affective reactions of students to-

wards educational activities, the learning setting, and the individuals within it), Group 2 exhibited a significantly stronger affective response ($p < 0.05$; $M = 3.69$) to the educational activities and learning environment than Group 1 ($M = 2.82$). This heightened emotional engagement suggests that the students in Group 2 developed a positive rapport with the AI feedback mechanism, potentially due to its perceived effectiveness or the novelty of the technology.

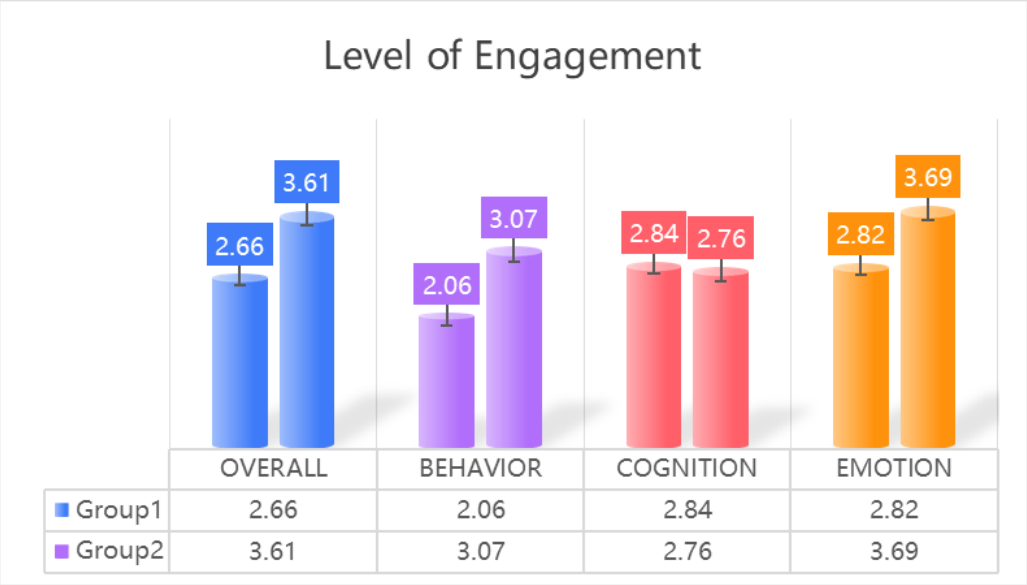

Figure 2. The level of engagement.

Table 5. The result of data analysis of academic engagement.

Test Statistics[a]				
	Overall	**Behavior**	**Cognition**	**Emotion**
Mann-Whitney U	9,242.500	9,898.000	17,731.500	11,143.000
Wilcoxon W	27,770.500	28,426.000	36,259.500	29,671.000
Z	−8.849	−8.075	−0.661	−6.923
Asymp. Sig. (2-tailed)	0.000	0.000	0.509	0.000

[a] Grouping Variable: TEST

Contrarily, to the Cognitive Engagement (assesses the focus and mental exertion that students direct towards the accomplishment of learning tasks), its scores did not reveal a significant difference between the two groups ($p = 0.509 > .05$), indicating that the level of mental exertion and focus directed towards learning tasks was comparable regardless of the feedback source. This finding implies that while the mode of feedback delivery (AI versus human) does not alter the cognitive effort exerted by students, it does impact their behavioral and emotional engagement.

The detailed analysis of these engagement metrics reveals a compelling trend: with appropriate instruction, doctoral students exhibit a pronounced inclination to engage more deeply with AI writing feedback, both in terms of the time dedicated and the emotional attachment. This preference for AI, particularly in the realms of behavioral and emotional engagement, may pose a significant challenge to the traditional role of teachers in academic writing support. As students become more adept at leveraging AI for writing improvement, there is a potential shift in the landscape of academic support, suggesting that educators and institutions may need to adapt their strategies to integrate AI effectively into the learning process [7, 29].

During the third phase of the study, a non-parametric independent samples test was utilized to explore five distinct facets distributed across three primary classifications of Academic Engagement. The objective of this analytical step was to delve deeper into doctoral students' perceptions regarding

the effectiveness of human versus AI instructors, specifically those in the humanities and social sciences. Such an investigation aims to provide critical insights that can inform the development of tailored instructional strategies for this cohort.

As shown in **Figure 3** and **Table 6**, the results elucidate a nuanced perspective on the comparative efficacy of writing feedback provided by teachers and AI systems. Specifically, the analysis reveals no statistically significant difference between the two feedback mechanisms in terms of Criterion-Based Comments ($p = 0.323 > 0.05$), suggesting that both AI and human educators adhere to established academic writing criteria as perceived by the doctoral students. This parity in adherence to criteria underscores the capability of AI to replicate the foundational aspects of academic feedback.

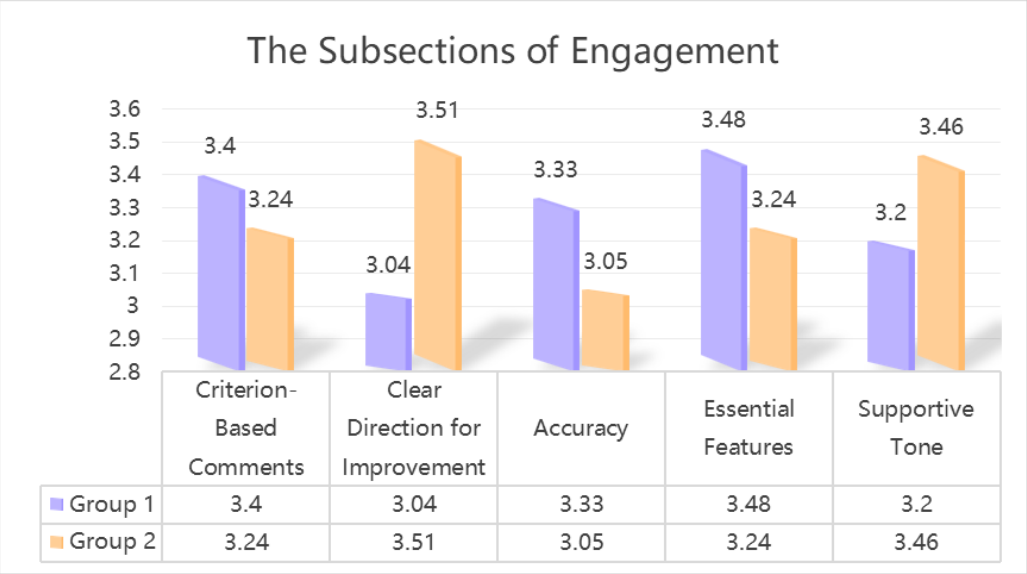

Figure 3. The subsections of engagement.

Table 6. The result of data analysis of the subsections of academic engagement.

Test Statistics[a,b]					
	CBC	**CDI**	**A**	**EF**	**ST**
Chi-Square	0.978	14.186	5.142	6.148	6.892
df	1	1	1	1	1
Asymp. Sig.	0.323	0.000	0.023	0.013	0.009

[a] Kruskal–Wallis Test
[b] Grouping Variable: TEST

However, a marked divergence is observed in other dimensions of feedback, such as the provision of Clear Direction for Improvement and the maintenance of a Supportive Tone. The AI feedback significantly outperforms that of teachers in these areas ($p = 0.000 < 0.05$, $M = 3.51 > 3.04$ for Clear Direction for Improvement; $p = 0.009 < 0.05$, $M = 3.46 > 3.20$ for Supportive Tone). The superiority of AI in providing explicit guidance for enhancement across various academic disciplines[5, 6], encompassing not only grammatical nuances but also complex areas such as data analysis and theoretical engagement, indicates a level of comprehensiveness that is challenging for human educators to match. The latter, constrained by the breadth of disciplines they must address and the limitations of their workload, often struggle to provide the same level of detailed and focused direction.

What is more, the AI's ability to maintain a Supportive Tone that aligns with instructional encouragement is particularly noteworthy. This finding challenges the conventional assumption that human interaction is inherently superior in

fostering a positive and encouraging learning environment. The AI's consistent and possibly algorithmically optimized supportive tone may serve as a valuable tool for bolstering student motivation and self-belief, an aspect where human educators, burdened by heavy workloads, may fall short[1, 4]. Conversely, in the realms of Accuracy and Essential Features of writing feedback, the feedback from teachers is found to be significantly more effective than that of AI ($p = 0.023 < 0.05$, $M = 3.33 > 3.05$ for Accuracy; $p = 0.013 < 0.05$, $M = 3.48 > 3.24$ for Essential Features). This indicates that while AI excels in breadth, it may lack the precision and discernment that human educators bring to the table. The tendency of AI to over-interpret and its occasional lack of moderation in issue identification can lead to inaccuracies, which are less prevalent in the more measured and nuanced feedback provided by teachers[29, 30]. Additionally, teachers' feedback tends to focus on the most fundamental and actionable improvement points, prioritizing the aspects that are most critical for student development.

5. Results and Discussion

To facilitate a more in-depth discussion of the comparative merits of writing feedback provided by AI and human instructors, **Figure 4** delineates the specific data details of each approach and presents their respective advantages and disadvantages more directly. This detailed breakdown provides a clearer basis for further analysis. **Figure 4** serves as a visual and analytical tool, juxtaposing the comparative efficacy of AI-generated feedback against that provided by human instructors in the context of English academic writing. Hoping that this detailed comparison not only illuminates the distinct strengths and weaknesses inherent in each modality but also provides clearer insights into the strategic potential of integrating these two approaches within pedagogical frameworks. The overarching goal of such an integrated model is to synergistically harness the unique advantages offered by both AI and human educators, thereby circumventing the limitations characteristic of each when employed in isolation.

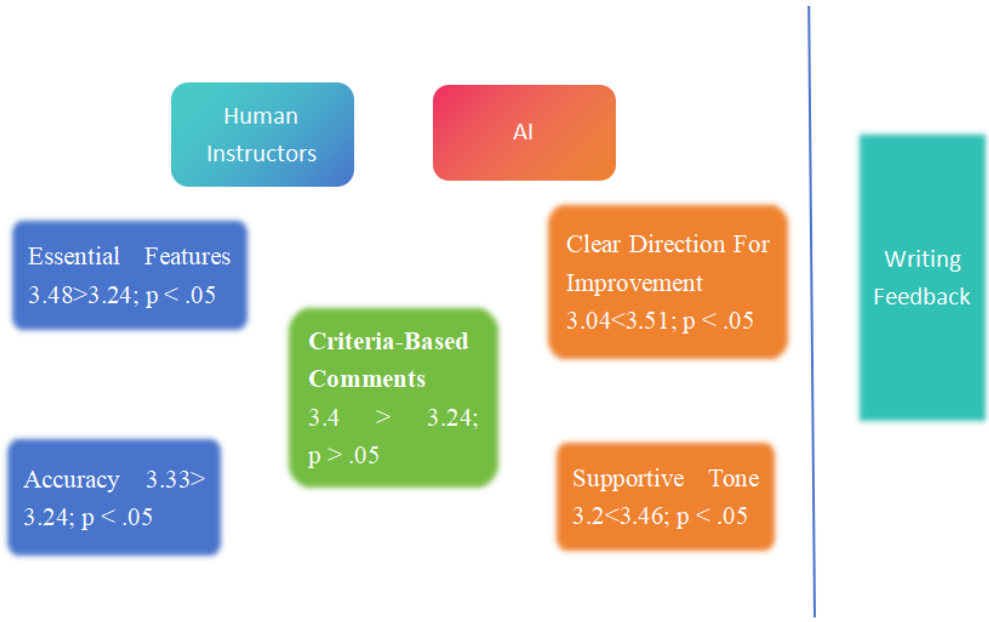

Figure 4. The comparison of writing feedback between AI and human instructors.

The empirical data presented in the analysis underscore the particular pertinence of AI's capacity for delivering explicit, comprehensive guidance, especially for doctoral students navigating the complex and intricate aspects of advanced academic writing. At the doctoral level, proficiency extends far beyond mere grammatical accuracy and stylistic elegance; it encompasses the sophisticated articulation of intricate arguments, the nuanced interpretation of complex data, and the critical engagement with theoretical paradigms. Studies by Deeva et al. (2021) and Drachsler (2023) have begun to elucidate how AI systems can be configured to provide feedback across these multifaceted domains, addressing

specific disciplinary conventions, offering targeted suggestions for refinement, and even identifying potential avenues for further scholarly inquiry[5, 6]. This capability to deliver domain-specific feedback across a broad spectrum of writing challenges represents a significant asset, potentially accelerating the developmental trajectory of doctoral candidates—a feat that can prove exceptionally challenging for human educators burdened by extensive responsibilities and finite time resources.

Furthermore, the finding that AI can consistently maintain a supportive tone, one that rivals or potentially surpasses the variability often observed in human interaction, acquires heightened significance within the context of the doctoral journey. This protracted and often arduous academic pursuit is frequently punctuated by periods of isolation, self-doubt, and fluctuating motivation, factors that can significantly impede progress. The consistent, algorithmically optimized encouragement provided by AI feedback can function as a vital psychological bulwark, mitigating feelings of isolation and furnishing a steady stream of positive reinforcement. This is particularly crucial given that support from human supervisors, while invaluable, can sometimes be inconsistent or inadvertently infrequent due to their heavy workloads and competing demands[1]. It is crucial to emphasize that this observation does not seek to diminish the irreplaceable value of human empathy and mentorship, but rather to acknowledge the supplementary role AI can play in reinforcing the emotional positivity essential for student perseverance.

However, the data, as further detailed in **Figure 4**, also reveals discernible limitations in AI performance, particularly concerning the accuracy of feedback and the precise identification of the essential features underpinning effective academic writing. This suggests that while AI demonstrates remarkable strengths in providing broad coverage and unwavering consistency, it currently lacks the intuitive depth, contextual discernment, and nuanced understanding inherent in human expertise. Instances where AI over-interprets textual elements or flags issues with insufficient nuance can lead to feedback that, rather than being pedagogically beneficial, introduces inaccuracies or distracts from genuinely critical areas for development. Conversely, human instructors leverage their experiential wisdom and pedagogical judgment to deliver feedback characterized not only by accuracy but also by a sophisticated tailoring to the individual student's specific needs, prior knowledge, and developmental trajectory. They possess the capacity to hierarchically prioritize feedback, concentrating on the most pivotal aspects requiring improvement, and to offer a richer, more contextually grounded appreciation of the essential hallmarks of rigorous academic prose.

Therefore, the synthesized analysis suggests that AI feedback holds considerable promise for augmenting the English academic writing competencies of doctoral students. Its capacity to deliver comprehensive, clear, and consistently supportive guidance can effectively address many of the pedagogical challenges encountered by this cohort. Nevertheless, it is imperative to recognize that AI, in its current state, is not a panacea for all instructional needs. The nuanced comprehension, contextual acuity, and critical judgment of human instructors remain indispensable for ensuring the feedback's ultimate accuracy, relevance, and developmental efficacy. The most efficacious pedagogical strategy likely resides in a carefully orchestrated blended model, wherein AI and human feedback operate in synergistic complementarity, each reinforcing the other's strengths to foster an optimal environment for maximizing the development of advanced English writing proficiency. Future research endeavors should prioritize exploring the most effective methodologies for integrating these two feedback modalities, with the ultimate aim of cultivating a truly synergistic, supportive, and intellectually enriching learning ecosystem tailored to the unique demands of doctoral education.

6. Conclusion

The findings of this study reveal a nuanced interplay between AI writing feedback and its human counterpart regarding Academic Engagement, particularly within the context of doctoral education in the humanities and social sciences. While AI writing feedback systems have undoubtedly demonstrated a remarkable capacity to provide comprehensive and supportive guidance[1, 4], often surpassing human instructors in terms of speed and volume of feedback, a critical analysis reveals certain limitations that warrant further consideration. Specifically, the analysis indicates that AI feedback, in its current iteration, tends to fall short in terms of quantifiable accuracy and a laser-like focus on the most essential features of academic writing pertinent to the specific disciplinary con-

text of the humanities and social sciences. These limitations potentially stem from the inherent challenges in training AI models to recognize and prioritize subtle nuances of argumentation, disciplinary-specific conventions, and the intricate interplay of ideas that characterize advanced scholarly work in terms of Academic Engagement among doctoral students of the humanities and social sciences.

These findings have significant implications for the pedagogical strategies employed in doctoral programs through Academic Engagement. They suggest that doctoral students would derive the most significant benefit from a symbiotic or "hybrid approach" to writing development[11], one that strategically leverages the complementary strengths of both AI and human educators. Such an approach would not simply treat AI and human feedback as interchangeable, but instead would recognize the unique affordances of each. Doctoral students should be actively encouraged to seek a deliberately balanced and diversified array of feedback. Specifically, they should be guided to utilize AI feedback tools for their breadth of coverage, their ability to identify surface-level errors and provide general writing support[8, 30], and their consistently "encouraging and supportive tone"[4]. Concurrently, students should be advised to rely on human instructors for their nuanced understanding of disciplinary-specific writing conventions, their ability to provide precise and targeted feedback on the depth and clarity of argumentation, and their expertise in identifying critical areas for improvement that align with the specific demands of the field. Furthermore, human educators can offer a level of mentoring and intellectual engagement that is currently beyond the capabilities of AI[9, 31].

In essence, by exploring the Academic Engagement of AI writing feedback and that of human instructors among doctoral students of the humanities and social sciences, this study advocates for a strategically integrated pedagogical model that views AI and human feedback as interdependent components of a comprehensive writing development framework regarding doctoral education in the humanities and social sciences. Future research should continue to investigate the evolving capabilities of AI writing feedback systems, while concurrently exploring practical pedagogical approaches for fostering a synergistic relationship between AI and human instructors in the development of advanced academic writing skills among doctoral students in the humanities and social sciences. Since this group is in urgent need of English academic writing improvement[11], compared to their counterparts in science and engineering, this ongoing investigation is crucial as we navigate the evolving landscape of higher education in doctoral students of the humanities and social sciences in the age of artificial intelligence.

Author Contributions

Methodology, Y.H.; investigation, X.L.; writing—original draft preparation, Y.H.; writing—review and editing, Y.H., X.L., and C.L.; supervision, C.L. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

The studies involving the participation of human beings were approved by BEIJING HUMANITIES AND SOCIAL SCIENCES ACADEMIC ETHICS COMMITTEE. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent (digital) to participate in this study. .

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

The data presented in the study can be inquired to the corresponding author.

Acknowledgments

We thank the students, teachers and the editors & reviewers who have contributed to this research, and the university concerned for the strong support.

Confict of Interest

The authors declare that the research was conducted without any potential conflict of interest.

References

[1] Steiss, J., Tate, T., Graham, S., et al., 2024. Comparing the quality of human and ChatGPT feedback of students' writing. Learning and Instruction. 91, 101894. DOI: https://doi.org/10.1016/j.learninstruc.2024.101894

[2] Applebee, A.N., Langer, J.A., 2011. "EJ" extra: a snapshot of writing instruction in middle schools and high schools. English Journal. 100(6), 14–27.

[3] Graham, S., 2019. Changing how writing is taught. Review of Research in Education. 43(1), 277–303.

[4] Escalante, J., Pack, A., Barrett, A., 2023. AI-generated feedback on writing: insights into efficacy and ENL student preference. International Journal of Educational Technology in Higher Education. 20, 55.

[5] Deeva, G., Bogdanova, D., Serral, E., et al., 2021. A review of automated feedback systems for learners: classification framework, challenges and opportunities. Computers and Education. 162, 104094. DOI: https://doi.org/10.1016/j.compedu.2020.104094

[6] Drachsler, H., 2023. Towards highly informative learning analytics. Open Universiteit: Heerlen, Netherlands. DOI: https://doi.org/10.25656/01:26787

[7] Drachsler, H., Kalz, M., 2016. The MOOC and learning analytics innovation cycle (MOLAC): a reflective summary of ongoing research and its challenges. Journal of Computer Assisted Learning. 32(3), 281–290. DOI: https://doi.org/10.1111/jcal.12135

[8] Pardo, A., Jovanovic, J., Dawson, S., et al., 2019. Using learning analytics to scale the provision of personalised feedback. British Journal of Educational Technology. 50(1), 128–138. DOI: https://doi.org/10.1111/bjet.12592

[9] Bond, M., Khosravi, H., De Laat, M., et al., 2024. A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. International Journal of Educational Technology in Higher Education. 21(4), 1–41. DOI: https://doi.org/10.1186/s41239-023-00436-z

[10] Wang, Y., Xue, L., 2024. Using AI-driven chatbots to foster Chinese EFL students' academic engagement: an intervention study. Computers in Human Behavior. 159, 108353.

[11] Zhang, L., Bi, J., Qin, X., 2021. The dilemma of English academic publication for non-native English speaking doctoral students: review and countermeasures. Foreign Language World. (03), 64–72.

[12] Reeve, J., 2012. A Self-Determination Theory Perspective on Student Engagement. In: Christenson, S.L., Reschly, A.L., Wylie, C. (eds.). Handbook of Research on Student Engagement. Springer: New York, United States. pp. 149–172.

[13] Sang, Y., Hiver, P., 2021. Engagement and Companion Constructs in Language Learning: Conceptualizing Learners' Involvement in the L2 Classroom. In: Hiver, P., Al-Hoorie, A.H., Mercer, S. (eds.). Student Engagement in the Language Classroom. Multilingual Matters: Bristol, UK. pp. 17–37.

[14] Derakhshan, A., Fathi, J., 2023. Grit and foreign language enjoyment as predictors of EFL learners' online engagement: The mediating role of online learning self-efficacy. The Asia-Pacific Education Researcher. 33(4), 759–769. DOI: https://doi.org/10.1007/s40299-023-00745-x

[15] Zare, J., Delavar, K.A., Derakhshan, A., et al., 2024. The relationship between self-regulated learning strategy use and task engagement. International Journal of Applied Linguistics. 34(3), 842–861. DOI: https://doi.org/10.1111/ijal.12535

[16] Zare, J., Derakhshan, A., Zhang, L.J., 2024. Investigating the relationship between metastrategy use and task engagement in an EFL context: a structural equation modeling approach. Innovation in Language Learning and Teaching. 19(2), 105–121. DOI: https://doi.org/10.1080/17501229.2024.2337710

[17] Graham, S., Hebert, M., Harris, K.R., 2015. Formative assessment and writing: a meta-analysis. The Elementary School Journal. 115(4), 523–547.

[18] Wilson, J., Czik, A., 2016. Automated essay evaluation software in English language arts classrooms: effects on teacher feedback, student motivation, and writing quality. Computers and Education. 100, 94–109.

[19] Hattie, J., Timperley, H., 2007. The power of feedback. Review of Educational Research. 77(1), 81–112.

[20] Motz, B., Canning, E., Green, D., et al., 2021. The influence of automated praise on behavior and performance. Technology, Mind, and Behavior. 2(3), 1–12.

[21] MacArthur, C.A., 2016. Instruction in Evaluation and Revision. In: MacArthur, C.A., Graham, S., Fitzgerald, J. (eds.). Handbook of Writing Research, 2nd ed. Guilford: New York, NY, USA. pp. 272–287.

[22] Roscoe, R.D., Varner, L.K., Crossley, S.A., et al., 2013. Developing pedagogically-guided algorithms for intelligent writing feedback. International Journal of Learning Technology. 8(4), 362–381.

[23] Panadero, E., Jonsson, A., Pinedo, L., et al., 2023. Effects of rubrics on academic performance, self-regulated learning, and self-efficacy: a meta-analytic review. Educational Psychology Review. 35(4), 113. DOI: https://doi.org/10.1007/s10648-023-09823-4

[24] Bai, L., Hu, G., 2017. In the face of fallible AWE feedback: how do students respond?

Educational Psychology. 37, 67–81. DOI: https://doi.org/10.1080/01443410.2016.1223275

[25] Grimes, D., Warschauer, M., 2010. Utility in a fallible tool: a multi-site case study of automated writing evaluation. The Journal of Technology, Learning and Assessment. 8(6), 4–43.

[26] Ranalli, J., 2018. Automated written corrective feedback: how well can students make use of it? Computer Assisted Language Learning. 31(7), 653–674. DOI: https://doi.org/10.1080/09588221.2018.1428994

[27] Zhu, M., Liu, O.L., Lee, H.S., 2020. The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing. Computers and Education. 143, 103668.

[28] Van Steendam, E., Rijlaarsdam, G., Sercu, L., et al., 2010. The effect of instruction type and dyadic or individual emulation on the quality of higher-order peer feedback in EFL. Learning and Instruction. 20(4), 316–327.

[29] Rüdian, S., Heuts, A., Pinkwart, N., 2020. Educational text summarizer: which sentences are worth asking for? Proceedings of The DELFI 2020 - The 18th Conference on Educational Technologies of the German Informatics Society; September 14–17, 2020; Bonn, Germany. pp. 277–288.

[30] Zawacki-Richter, O., Marín, V.I., Bond, M., et al., 2019. Systematic review of research on artificial intelligence applications in higher education–where are the educators? International Journal of Educational Technology in Higher Education. 16(1), 1–27. DOI: https://doi.org/10.1186/s41239-019-0171-0

[31] Moore, N.S., MacArthur, C.A., 2016. Student use of automated essay evaluation technology during revision. Journal of Writing Research. 8(1), 149–175.