

Evaluating Advanced EFL Learners' Writing Performance and Perceptions of Large Language Model Assistance

Mengkai Wang* 

Beijing Foreign Studies University, China

Abstract

Since the release of Large Language Models (LLMs), they have been widely used in language teaching and acquisition. Previous studies have explored LLMs' assistance in improving language performance, but their impact on advanced English Foreign Language learners' writing remains unclear. This study conducts an experiment based on their IELTS or TOEFL scores. After a pre-test to exclude participants who could not follow instructions, they were randomly divided into two groups. Participants in the experiment group were instructed to complete the essay with guidance but not to directly generate it. Following the task, participants completed a motivation and perception questionnaire. Surprisingly, results show that LLMs did not significantly improve writing performance in advanced EFL learners. However, the questionnaire revealed strong positive perceptions regarding LLM utility and motivation. Additionally, it appears that the current level of LLMs matches the writing capabilities of advanced learners. This study explores how LLMs affect EFL performance and learner perceptions, offering insights into their potential assistance.

Keywords: EFL learners, writing performance, LLMs, learner perception, motivation

Introduction

The rapid integration of large language models (LLMs) such as ChatGPT, into educational contexts has generated widespread interest regarding their utility

*Corresponding Author E-mail: tayloricy@bfsu.edu.cn

for second language acquisition. However, the effectiveness of LLM assistance for advanced English as a Foreign Language (EFL) learners remains relatively underexplored. In line with the conference theme of “Inclusive CALL”, ensuring that educational technologies are truly inclusive requires understanding their utility across the entire spectrum of learner proficiency. Advanced learners, a group often overlooked in the design of support tools, represent a critical demographic for this investigation.

Therefore, this study addresses this research gap through a dual-focus approach. It primarily examines whether ChatGPT-assisted writing significantly improves the objective writing performance of advanced Chinese EFL learners. Concurrently, it explores learners’ subjective perceptions of LLM utility and motivation, aiming to bridge the potential gap between perceived usefulness and measurable performance outcomes.

Literature Review

Previous research predominantly focuses on intermediate-level learners, suggesting notable benefits from LLM assistance in grammar accuracy, lexical development, and overall writing proficiency (Barrot, 2023; Song & Song, 2023). These learners often find LLM-generated feedback and suggestions highly valuable for overcoming specific linguistic hurdles.

However, the context for advanced learners is theoretically different. Characterized by high proficiency levels, they typically operate near the upper bounds of their Zone of Proximal Development (ZPD), as conceptualized by sociocultural theory (Vygotsky, 1978). In this framework, learning is most effective when external support aids a learner in tasks slightly beyond their independent capability. For advanced learners who already possess strong autonomous writing skills, the scaffolding provided by current LLMs might offer only limited benefits (Liu et al., 2023). Consequently, they might derive less significant cognitive advantages from LLMs compared to their less proficient counterparts. This highlights a potential bifurcation in LLM impact: while cognitive gains may be minimal for this group, the affective benefits, such as increased motivation or reduced writing anxiety, could still be substantial.

Methodology

A randomized controlled design was employed, involving 50 advanced Chinese EFL learners recruited from university English programs. All participants met the advanced proficiency criteria (IELTS score ≥ 7.5 or TOEFL score ≥ 105).

Prior to the task, all participants in the experimental group received a 15-minute standardized training session. This session demonstrated how to use ChatGPT for brainstorming, outlining, and receiving feedback on grammar and vocabulary, while explicitly instructing them not to use it for generating complete sentences or paragraphs.

Participants were randomly assigned into two groups: Experiment Group and Control Group. Both groups completed a 1-hour argumentative essay task on contemporary topics. Essays were evaluated by three raters, an AI scoring engine, a trained researcher, and a certified IELTS instructor, according to IELTS assessment criteria, including fluency, grammatical accuracy, and lexical and grammatical complexity. In addition to writing performance assessments, participants in the experimental group completed a post-task questionnaire measuring motivation and perceptions of LLM utility. The questionnaire responses were analyzed using principal component analysis (PCA) to uncover underlying dimensions of learners' attitudes toward LLMs.

Results

Descriptive statistics suggested only minimal differences between groups, as shown in Table 1. Although the experimental group showed slightly higher scores in grammar and language use, independent samples t-tests revealed no statistically significant differences in any dimension ($p > .05$). Similarly, ANCOVA, controlling for pre-task self-rated writing confidence, confirmed the absence of significant group effects ($p = .35$). PCA results for the post-task questionnaire identified three principal components: non-motivation, motivation, and perceived assistance. Figure 1 presents a summary of the relationships among questionnaire items and components. Participants reported positive perceptions of LLM assistance. Mean ratings were high for motivation and assistance, suggesting learners felt that ChatGPT boosted their confidence and reduced writing anxiety. Nevertheless, no correlation was found between these positive perceptions and actual performance improvements.

Discussion

The findings of this study align with sociocultural theory predictions concerning the Zone of Proximal Development. Since advanced learners already possess strong autonomous writing skills, the scaffolding provided by ChatGPT offered minimal additional advantage. This lack of significant performance improvement contrasts with studies focusing on intermediate-level learners, where LLMs have been shown to yield more pronounced gains (Barrot, 2023;

Table 1. Descriptive statistics for writing outcomes

Rating Aspect	Descriptive Statistics			
	Group	Mean	Std. D	Std. E
Overall	Experiment	92.21	2.28	0.46
	Control	91.47	2.39	0.48
Content	Experiment	22.79	1.30	0.26
	Control	22.87	1.08	0.22
Grammar	Experiment	22.97	1.15	0.23
	Control	22.82	1.32	0.26
Language Use	Experiment	23.26	1.13	0.23
	Control	22.85	1.32	0.26
Coherence and Cohesion	Experiment	23.20	0.97	0.19
	Control	22.94	1.20	0.24

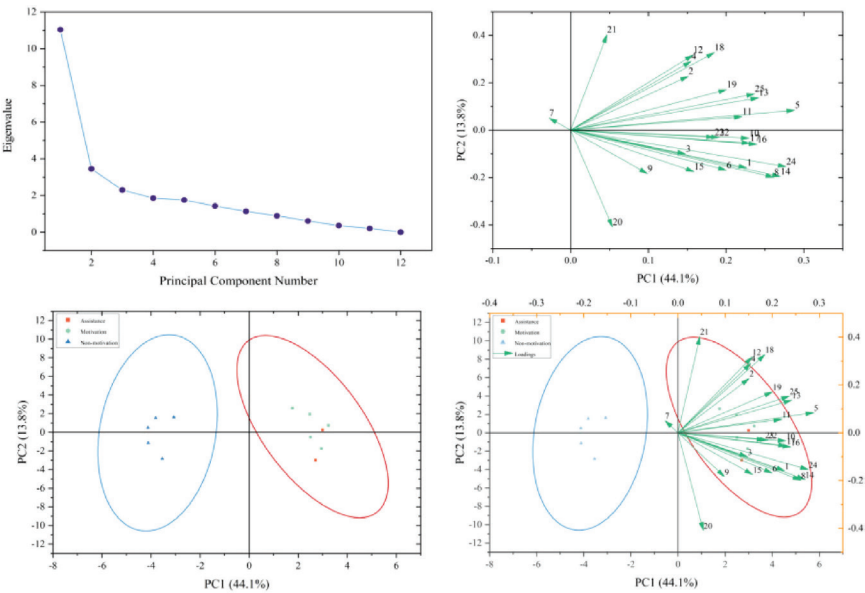


Figure 1. Visualized principal component analysis results.

Song & Song, 2023). Our results suggest that the “one-size-fits-all” application of LLMs may not be effective across different proficiency levels.

While real-time feedback and language suggestions were appreciated, these tools may not have sufficiently challenged learners to catalyze measurable growth. However, the positive motivational effects reported by participants

resonate strongly with earlier studies (Hwang et al., 2023; Song & Song, 2023), indicating that LLMs can enhance learners' emotional engagement even when cognitive gains are limited. This suggests a bifurcated role of AI in language education: improving affective variables without necessarily driving immediate skill acquisition for highly proficient learners.

Moreover, the slight grammatical and lexical advantages observed in the experiment group, although not statistically significant, hint at a possible cumulative benefit over longer durations. Given the short-term nature of this study, it remains plausible that extended interaction with LLMs could yield meaningful improvements for advanced learners. Importantly, the disconnect between learners' perceived benefits and actual performance outcomes raises critical pedagogical questions. While high satisfaction and motivation are valuable, educators must balance these factors with realistic expectations about LLMs' capacity to foster measurable writing improvements at higher proficiency levels. To be truly inclusive, CALL tools must be designed or implemented in a proficiency-sensitive manner.

Several limitations of the current study warrant attention. The homogeneous sample (Chinese advanced EFL learners) limits generalizability across cultural and linguistic backgrounds. Furthermore, the short intervention period may not capture the longitudinal effects of sustained LLM usage.

Conclusion

This study contributes to the emerging literature on AI applications in second language writing by demonstrating that, for advanced EFL learners, ChatGPT-assisted writing primarily enhances motivational factors rather than producing significant immediate improvements in writing performance. These findings highlight the need for proficiency-sensitive design of AI educational tools and call for continued exploration of the long-term impacts of AI integration in language learning. By focusing on a frequently overlooked learner group, this research underscores the importance of a nuanced, proficiency-aware approach to creating truly inclusive CALL environments.

References

- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, Article 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Liu, M., Zhang, L. J., & Biebricher, C. (2023). Investigating students' cognitive processes in generative AI-assisted digital multimodal composing

- and traditional writing. *Computers & Education*, 211, Article 104977. <https://doi.org/10.1016/j.compedu.2023.104977>
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, Article 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.