



Digital Applied Linguistics, 1, 102324 (2024)
<https://doi.org/10.29140/dal.v1.102324>



Examining the effect of generative AI on students' motivation and writing self-efficacy

JERRY HUANG^a (Corresponding Author)  

ATSUSHI MIZUMOTO^a 

^a Kansai University, Faculty of Foreign Language Studies, Japan

Abstract

The present study explores the effects of generative AI, specifically ChatGPT, in EFL classrooms on student motivation and writing efficacy. Motivation was measured through three components: Ideal L2 Self (IL2), Ought-to L2 Self (OL2), and L2 Learning Experience (L2LE). Participants ($n = 327$) were first and second-year undergraduate students at a Japanese university, enrolled in mandatory English classes focused on reading/writing or speaking/listening. The control group ($n = 164$) received peer feedback, whereas the treatment group ($n = 163$) utilized ChatGPT with specially crafted prompts for feedback. Both groups completed pre- and post-questionnaires to assess motivation and writing self-efficacy. Results affirmed that ChatGPT positively affected students' motivation related to Ideal L2 Self and L2 Learning Experience. ChatGPT also significantly enhanced writing self-efficacy, which was found to correlate with all three motivational factors. However, there was no impact on Ought-to L2 Self motivation. The study highlights that ChatGPT's integration can improve intrinsic motivation and writing self-efficacy, provided structured guidance is available to manage issues such as plagiarism. Future research should examine diverse samples, long-term effects, and ChatGPT's impact on other language skills.

Keywords: The L2 motivational self system; ideal L2 self; ought-to-L2 self; L2 learning experience; generative AI; ChatGPT; self-efficacy

Introduction

In the field of digital applied linguistics, the rapid advancement of Artificial Intelligence (AI) technology has resulted in significant transformations, creating a growing demand for the integration of AI into writing curriculum development (Hao et al., 2024; Yang et al., 2024). AI technologies can provide personalized feedback, facilitate collaborative writing experiences,

and offer engaging writing prompts, all of which may foster a more motivating learning environment. Motivation has long been recognized as a critical factor in applied linguistics, influencing learners' engagement, persistence, and success (Dörnyei, 2020). With the advent of digital tools and platforms, it becomes increasingly essential to explore how AI technology can enhance or hinder motivation among language learners. One of the latest technological advancements in AI technology is the introduction of AI chatbots, such as ChatGPT, released in late 2020. These chatbots offer interactive and personalized language practice opportunities, which can potentially boost motivation by providing immediate feedback, simulating real-life conversations, and accommodating individual learning paces. Despite the promising potential of AI chatbots, the specific role of ChatGPT feedback in learners' motivation remains largely under-researched.

A recent study conducted in the United States (McDonald et al., 2024) surveyed 116 top research (R1) universities. Findings show that only a modest number (63%) actively encourage the use of generative AI (GenAI), and roughly half offer extensive classroom guidance, sample syllabi, and curriculum support, primarily centered on writing activities. However, the widespread availability of smartphones allows learners easy access to information and tools, such as computer-assisted language learning, both inside and outside the classroom. Common resources such as Google Translate and DeepL are already widely used among learners, and studies (e.g., Stapleton & Leung, 2019) have thoroughly examined their effectiveness. With the rise of GenAI, learners now have access to a broader range of tools that can enhance their language learning experience. Recent research indicates that ChatGPT may foster learner autonomy (Augustini, 2023), enhance students' motivation and writing skills (Song & Song, 2023), and encourage pleasure-based motivation (Cai et al., 2023).

Although the use of ChatGPT appears to positively influence motivation, the commonly referenced L2 Motivational Self-System (L2MSS) in second language acquisition (SLA) has rarely been applied in this context. Huang and Mizumoto (2024a) demonstrated ChatGPT's effectiveness in maintaining student motivation. The present study builds upon their work by expanding the participant pool and adding an analysis of learners' perceptions of their self-efficacy in writing after using ChatGPT.

This study seeks to provide a more comprehensive understanding of how ChatGPT influences learners' motivation. By expanding the participant pool and incorporating an analysis of learners' perceptions of their self-efficacy in writing after using ChatGPT, it offers a deeper examination of the alignment between ChatGPT's impact and the principles of the L2MSS. Additionally, the study explores how learners' motivational self-systems interact with AI-driven language learning environments.

Literature Review

L2 Motivational Self System

Following Zoltán Dörnyei (2005), the L2 Motivational Self System (L2MSS) has become a major focus of study, spurring substantial research on individual differences in language learning motivation (Al-Hoorie, 2018). Although questions have been raised about the instrument's validity (Al-Hoorie et al., 2023), other researchers have reaffirmed the system's reliability (Henry & Liu, 2024; Papi & Teimouri, 2024). Despite recent debates, this framework remains a significant development in SLA as it explores the psychological factors that drive learners' motivation to study a second language (L2). Comprising three components—the Ideal L2 Self, the Ought-to L2 Self, and the L2 Learning Experience—the L2MSS examines how different aspects of learners' self-perception influence motivation. The Ideal L2 Self, seen as the strongest motivating factor, represents learners' envisioned future self as a fluent language

user. The Ought-to L2 Self reflects external pressures and expectations about language proficiency, while the L2 Learning Experience encompasses learners' real interactions in the language, which can affect motivation either positively or negatively.

Dörnyei (2005) describes the Ideal L2 Self as the specific part of a learner's ideal identity tied to their proficiency in an L2, embodying the inspiring vision of becoming fluent and socially capable in the language. This aspirational self-image strongly motivates individuals to narrow the gap between their current skills (actual self) and the ideal self they aim to achieve. Studies by Taguchi et al. (2009) and Ryan (2009) have found that this concept is closely linked to integrativeness and accounts for a significant portion of the variance in learners' effort.

The Ought-to L2 Self reflects learners' sense of duty or obligation to learn an L2, shaped primarily by external expectations. It represents qualities learners believe they should develop based on responsibilities or others' expectations. Taguchi et al. (2009) found in a comparative study across Japan, China, and Iran that family influence and preventive motivations were key factors in shaping this aspect. Further research in Hungary by Csizér and Kormos (2009) noted a positive link between parental support and the Ought-to L2 Self.

The L2 Learning Experience refers to how learners feel about language learning, often influenced by immediate factors such as the classroom environment. According to Dörnyei (2009), many learners find motivation in successful experiences within the learning process itself rather than through pre-existing or externally imposed self-images. This aspect is affected by factors like the curriculum, the L2 instructor, classmates, and teaching materials.

Generative AI and Writing

The launch of OpenAI's ChatGPT in November 2022 has revolutionized academic writing, showcasing its ability to produce "high-quality" papers suitable for journal submissions. Teng (2023) noted that with ongoing AI advancements, significant changes in academic publishing are anticipated, suggesting that the field is still in the early stages of transformation. However, while large language models present exciting potential for enhancing education and supporting instructors, their limitations and potential biases must be carefully considered to ensure responsible usage (Kasneci et al., 2023). Yang et al. (2024) studied the integration of ChatGPT in an EFL academic writing course, finding it beneficial for tasks like proofreading, brainstorming, and translation. However, the study stresses the importance of using ChatGPT thoughtfully, framing it as a "ghostwriter" that should be navigated with caution. Prioritizing learners' development as critical thinkers and emerging scholars remains essential to avoid over-reliance on AI and to foster genuine academic growth.

This new technology has generally been met with enthusiasm. Tseng and Warschauer (2023) introduced a five-part educational framework to help L2 learners engage with AI by focusing on understanding, accessing, prompting, verifying, and integrating AI effectively. By teaching students how to interact with AI, educators can prepare them for technology's evolving role both in and out of the classroom. Additionally, Cotton et al. (2023) underscores the need for clear communication about assessment standards and academic integrity to minimize issues such as cheating in university settings. As these pedagogical challenges emerge, Hao et al. (2024) advocate for a balanced approach to AI integration within English major programs, emphasizing the importance of developing both AI proficiency and critical thinking skills to prepare students for a technology-driven workforce. The authors propose a theoretical and pedagogical model that supports AI and critical thinking integration through a dynamic learning environment, curriculum evolution, research enhancement, and faculty and student support.

Research on ChatGPT is in its early phases but reveals significant potential for educational development, particularly in L2 learning. ChatGPT has been shown to enhance learner

perceptions (Marzuki et al., 2023; Xiao & Zhi, 2023; Van Horn, 2024), boost writing and language skills (Athanasopoulos et al., 2023; Song & Song, 2023), and support teachers in L2 instruction (e.g., Barret et al., 2023; Jeon & Lee, 2023; Mohamed, 2024). ChatGPT's integration into the writing process, from planning to editing, offers structured support while maintaining integrity (Barrot, 2023). Huang (2023) demonstrated effective pedagogical practice by using well-designed prompts to help students receive personalized feedback on their work. These findings underscore ChatGPT's growing role as a valuable tool for enhancing both teaching and learning experiences in L2 education.

As more studies emerge, the technology's role in education is likely to expand, setting new standards for instructional methods. For example, Kohnke et al. (2023) found ChatGPT useful for enhancing students' comprehension and assisting teachers with lesson planning. Guo and Wang (2023) suggest that teachers incorporate ChatGPT's essay feedback into their own responses to students. This technology is positioned to make a lasting impact, prompting questions about its effects on L2 learner motivation and identity.

There were some studies specifically on ChatGPT and L2 writing. Research into the use of ChatGPT for writing assessment has shown promising results. Teng (2024b) explores the role of ChatGPT in EFL writing through a systematic review of the 20 most relevant articles, following PRISMA guidelines. The findings indicated that while ChatGPT provides opportunities for enhancing writing skills—such as instant feedback and diverse prompts—it also presents challenges, including potential dependency on AI and the necessity for critical thinking, thus advocating for a balanced approach to integrating AI tools like ChatGPT into writing curricula and fostering a community of practice among educators and students as recommendation. Mizumoto and Eguchi (2023) investigated its application in L2 writing environments, specifically focusing on Automated Essay Scoring (AES). By analyzing a dataset of 12,100 essays, they found that AES systems powered by GPT technology demonstrated significant accuracy and reliability, closely matching human evaluations. This suggests that ChatGPT can provide automated corrective feedback (CF), making it a valuable resource for writing instructors in educational settings. Building on this, Mizumoto et al. (2024) assessed ChatGPT's effectiveness in evaluating accuracy in L2 writing. Using the Cambridge Learner Corpus First Certificate in English (CLC FCE) dataset, they compared ChatGPT's performance with human evaluators and Grammarly. The study revealed a high correlation between ChatGPT's assessments and human ratings, with a correlation coefficient (ρ) of 0.79, surpassing Grammarly's 0.69. This highlights ChatGPT's precision in automated evaluations, confirming its suitability as a tool for L2 writing assessment. In another study, Allen and Mizumoto (2024) examined the experiences of 33 Japanese EFL learners who used writing groups alongside AI technology for editing and proofreading. The students expressed a preference for AI tools, like ChatGPT, for editing and proofreading, appreciating the effective feedback that improved the clarity and cohesion of their writing. ChatGPT provided authoritative insights and feedback, similar to peer review processes in writing groups. The authors recommended integrating ChatGPT with writing groups to enhance editing and proofreading practices. Additionally, Teng (2024c) explored the impact of ChatGPT on 45 EFL learners in Macau, focusing on their perceptions and experiences with AI-generated feedback. The findings highlighted significant positive effects of AI assistance on writing, including increased motivation, self-efficacy, engagement, and a tendency towards collaborative writing. Qualitative analysis identified four key themes: the influence of AI assistance on writing self-efficacy, engagement, motivation, and the potential for collaborative writing. These studies collectively underscore the transformative potential of ChatGPT in enhancing writing skills and educational practices.

Generative AI Use and Motivation

The use of GenAI in EFL classrooms is still emerging, but research has already shown a positive connection between AI and student motivation. For instance, a study by Leong et al. (2024) investigated whether GenAI could personalize vocabulary-learning examples to enhance both learning outcomes and motivation. Researchers developed a web app with three conditions: a control group using pre-existing example sentences, and two experimental groups where participants generated sentences or short stories based on their interests. Although no significant difference in learning outcomes was found across the conditions, participants in the AI-supported groups reported higher intrinsic motivation, a stronger sense of choice, and greater feelings of competence compared to the control group. This suggests that while AI personalization may not directly affect learning outcomes, it can positively influence students' motivation. Another study by Wei (2023), conducted at a university in China, found that AI-mediated language instruction using the Duolingo platform significantly improved English learning achievement, L2 motivation, and self-regulated learning among EFL learners, compared to traditional methods. This was further supported by qualitative interviews with students in the experimental group, who viewed the AI platform as engaging, personalized, and empowering, which led to increased motivation, confidence, and a more positive learning environment. Additionally, Yamaoka (2024) found that ChatGPT acted as a personalized teaching assistant, boosting Japanese EFL learners' motivation and reducing their anxiety by allowing them to take control of their learning. However, some learners experienced negative motivational impacts, such as anxiety over potential misinformation or a loss of confidence due to ChatGPT's high-quality proofreading. Together, these studies highlight the positive impact of GenAI on learning motivation, despite potential challenges.

The studies above focus on the influence of AI on student motivation, rather than the reverse. A study by Zheng et al. (2024) investigated the technology acceptance of ChatGPT among EFL learners, specifically examining the factors outlined in the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) model, as well as the moderating role of motivation as conceptualized by Self-Determination Theory (SDT). The results showed that performance expectancy, effort expectancy, social influence, hedonic motivation, habit, and SDT motivation were significant predictors of learners' behavioral intention to use GenAI tools. Additionally, the study found that SDT motivation significantly moderated the relationships between certain UTAUT2 constructs and learners' behavioral intention and actual use of GenAI tools, highlighting the important role of motivation in shaping technology acceptance. Lai et al. (2023) explored the factors influencing Hong Kong undergraduate students' intentions (as defined by behavioral intent in the Technology Acceptance Model) to use ChatGPT to support active learning. The study found that intrinsic motivation and perceived usefulness were the strongest predictors of students' intention to use ChatGPT for answering academic inquiries. Similarly, Huang and Mizumoto (2024b) found that in Japanese EFL classrooms where ChatGPT is a required learning tool, Ought-to L2 Self motivation became a significant predictor of technology acceptance, specifically influencing students' actual use of ChatGPT. This suggests that when students feel a sense of obligation or external pressure to learn English, they are more likely to adopt new technologies to meet these expectations. However, the study emphasized the need to explore additional components within the L2MSS that might yield similar effects. These studies further reinforce the connection between GenAI and motivation.

Feedback and Motivation

Feedback and motivation are closely related. Fang (2023) examined how written feedback impacts writing motivation and self-efficacy among Chinese EFL high school students. The findings showed that while feedback significantly increased students' motivation to write, it did not substantially influence their self-efficacy in writing. This suggests that feedback can encourage students to engage more in writing but may not necessarily improve their confidence in their writing abilities. Similarly, Ahmetovic et al. (2023) explored the role of CF in motivation and EFL achievement among middle and high school students in Bosnia and Herzegovina. The study revealed that students generally appreciate CF and view it positively, believing it enhances their learning. While a positive attitude toward CF was linked to increased motivation for speaking and writing in English, neither oral nor written CF significantly predicted students' overall EFL achievement. A meta-analysis by Cen and Zheng (2023) offers further insight into the relationship between feedback and motivation. The study highlighted several important points: when delivered effectively, feedback can be a powerful motivator; personalized feedback is essential, as its impact varies among students; students who view feedback positively are generally more motivated; feedback encourages self-regulation, which indirectly boosts motivation; and a supportive classroom environment can enhance the motivational benefits of feedback. These findings suggest that GenAI could contribute meaningfully by providing personalized feedback, reducing anxiety, promoting positive attitudes, and fostering a supportive context. Thus, feedback from GenAI could indirectly enhance students' motivation, aligning with the patterns observed in previous studies.

Motivation and Writing Self-Efficacy

Self-efficacy, a central concept in social cognitive theory, refers to individuals' beliefs in their ability to organize and execute actions to reach specific goals (Bandura, 1997). These beliefs have a strong influence on performance, often predicting outcomes more reliably than actual skill level (Bandura, 1997; Schunk, 1991). Research indicates that self-efficacy is not only a crucial motivational factor but also shapes students' choices, effort, persistence, thought processes, and emotional responses to writing tasks (Zhang et al., 2023). In an investigation of online English learning among Chinese university students, Teng (2024a) explored the relationships between social support, self-efficacy, emotional adjustment, and foreign language anxiety (FLA). The findings show that social support significantly predicts emotional adjustment and self-efficacy beliefs, and that self-efficacy acts as a mediator between social support and FLA. This suggests that fostering self-efficacy through social support may be an effective way to alleviate FLA and improve outcomes in online English learning. Additionally, a study by Teng and Yang (2022) examined the interplay among metacognition, motivation, self-efficacy beliefs, and English learning outcomes in online learning environments. Their results revealed that self-efficacy significantly predicts online English learning achievement, both directly and indirectly, through motivation and metacognition. This study underscores the value of a supportive online learning environment that cultivates self-efficacy, motivation, and metacognitive awareness to enhance learners' experiences. Given these insights, the introduction of GenAI in EFL settings may also contribute positively. As Huang and Mizumoto (2024a) highlight, increased motivation could, in turn, lead to enhanced self-efficacy in writing, showing the potential for GenAI to support and elevate learning in EFL contexts.

Research Questions

Previous research showed a slight increase in all three motivation factors, though the increase in the Ought-to L2 self was not statistically significant. This study aims to replicate those findings while also comparing students' self-reported writing efficacy before and after using GenAI.

Additionally, it also seeks to investigate the relationship between motivation and self-efficacy following GenAI usage. Based on the literature review, the research questions are as follows:

- RQ1:** Do students report higher Ideal L2 Self motivation after using GenAI in the language classroom?
- RQ2:** Do students report higher Ought-to L2 motivation after using GenAI in the language classroom?
- RQ3:** Do students report higher L2 Learning Experience motivation after using GenAI in the language classroom?
- RQ4:** Do students report higher writing self-efficacy after using GenAI in the language classroom?
- RQ5:** Do the correlations between writing self-efficacy and motivational factors increase after using GenAI in the language classroom?

Materials and Methods

Participants

This study involved 327 first- and second-year undergraduate students taking a required English course. Of these, 164 students in a Listening/Speaking class formed the control group, while 163 students in a Reading/Writing class made up the treatment group. In the Listening/Speaking class, students individually prepared scripts for group presentations and received feedback from their peers. Meanwhile, in the Reading/Writing class, students collaborated on essays, with each student responsible for a different paragraph, such as the introduction, body, or conclusion. Both groups completed their respective tasks twice over the course of the semester.

Instruments and Data Collection

Students completed a survey twice during the semester—once at the beginning and once at the end—to capture potential changes in motivation and writing self-efficacy over time. The study used a 22-item questionnaire (see Appendix), adapted from the works of Taguchi et al. (2009) and Pajares and Valiante (1999). Taguchi et al.'s (2009) L2 Motivational Self System (L2MSS) questionnaire, initially developed for Japanese participants, provided the framework for assessing three core motivational components: Ideal L2 Self (IL2), Ought-to L2 Self (OL2), and L2 Learning Experience (L2LE).

- IL2 items prompted students to envision themselves confidently using English in their future, with questions like, “I can imagine myself living abroad and having a discussion in English.”
- OL2 items reflected external pressures or expectations, such as “I study English because close friends of mine think it is important.”
- L2LE items focused on students' current engagement in English classes, with items like “I really enjoy learning English.”

In addition, ten questions were adapted from Pajares and Valiante's (1999) writing self-efficacy scale to assess students' confidence in specific writing skills, such as “I can structure paragraphs to support ideas in the topic sentences” and “I can write simple sentences with good grammar.”

To ensure the integrity of responses and minimize potential biases, the university's online learning management system randomized the sequence of questions for each participant. Any

incomplete survey prompted a reminder for students to finish. All survey questions were presented in Japanese, and responses were recorded on a 6-point Likert scale, with options ranging from 1 ("strongly disagree") to 6 ("strongly agree").

Comparison of Control and Treatment Groups

Participants in this study were students from a private university in Japan's Kansai region, enrolled in either a required English reading and writing class or a speaking and listening class. While individual English proficiency was not directly measured, it was estimated to range from A2 to B2 on the CEFR scale, based on guidelines from Aizawa et al. (2020). All participants had completed at least eight years of English instruction as part of Japan's mandatory education curriculum in primary and secondary schools.

Students in the control group participated in speaking and listening classes, which included two group oral presentations during the semester. Presentation topics were aligned with the university's prescribed textbook to maintain consistency. Prior to each presentation, students used an entire class period to develop their speaking scripts collaboratively, receiving peer feedback in groups of four or five. This feedback was intended to improve transitions and script coherence, moving beyond basic comments. Based on their peers' input, students revised their scripts, completing them either in class or as homework. All tasks were done on paper, and final scripts were submitted as part of the assessment after the oral presentations.

The treatment group consisted of students in reading and writing classes, where they collaborated on two group essays over the semester. Essay topics were similarly aligned with the textbook to support writing skills development. Each group of five students worked on one part of the essay, with each student responsible for a specific paragraph (e.g., introduction, body, or conclusion). On the first day of the writing process, students created a flowchart, outlined, and drafted their essays on paper to limit reliance on ChatGPT. Any incomplete work could be finished at home. On the second day, students attended a brief lecture on using ChatGPT, with pre-made prompts focused on providing structured feedback on essay organization, coherence, and structure (Huang, 2023). Students received training on plagiarism prevention and were required to avoid copy-pasting text. Any revisions based on AI feedback were documented on paper and submitted through the university's online learning system for further review. By the end of the second week, students submitted the final version of their essays along with paper drafts showing their revision history.

Data Analysis

The data for this study was processed using R software (version 4.3.3). To answer the first four research questions, the dataset was analyzed for normality through Shapiro-Wilk p -values (w) and for homogeneity of variance using standard deviations. Instrument reliability was assessed by calculating Cronbach's alpha (α). Interaction effects between factors were explored using a two-way analysis of variance (ANOVA), which was appropriate due to the dataset meeting the assumptions of normality, homogeneity, and reliability. This approach enabled the comparison of mean scores between two groups and assessed the statistical significance of post-intervention differences. To address the fifth research question, Pearson's correlation coefficient (r) was employed to determine the nature and strength of the relationship between motivation and self-efficacy in both groups before and after the intervention. This analysis was essential for evaluating how the intervention influenced these psychological constructs.

To ensure the reproducibility and transparency of the data analysis process, the data and the R code used in the study have been made accessible on OSF (<https://osf.io/h4up2>).

Results

The following descriptive statistics provide insights into the effects of treatment and time on three L2MSS constructs and writing self-efficacy. Table 1 shows that the mean scores for these variables range from 3.36 to 4.10, with standard deviations between 0.81 and 1.34, indicating relatively consistent variance across the groups. Shapiro-Wilk test *p*-values, ranging from 0.97 to 0.99, suggest a normal distribution for all variables within each group. Additionally, Cronbach's alpha (α) values, which range from 0.75 to 0.92, demonstrate strong internal consistency and reliability of the variables measured. The treatment group's mean scores show notable increases across all four variables (3.55 to 3.71, 3.42 to 3.61, 4.29 to 4.47, and 3.52 to 3.95), whereas the control group's mean scores remained relatively stable with slight variability (3.55 to 3.55, 3.36 to 3.43, 4.24 to 4.22, and 3.63 to 4.10). Given the normal distribution, homogeneity, and reliability of the data, ANOVA can be used to test for significant differences and interaction effects, assessing whether the post-intervention score increases in these variables are statistically significant.

A two-way ANOVA was performed to examine the interaction between ChatGPT usage and pre-post scores for each variable. The findings are presented in Table 2.

The results of Ideal L2 self in Table 2 shows no significant effect of treatment alone ($p = .49$) with a negligible effect size (Partial $\eta^2 < .01$). However, time had a significant effect ($p < .05$) with a small effect size (Partial $\eta^2 = .01$), indicating that changes over time are meaningful.

Table 1 Descriptive statistics.

Variables	Group	Test	<i>n</i>	Mean	<i>SD</i>	<i>p</i>	α
Ideal L2 Self	Treatment	Pre	163	3.55	1.03	Pre .98	Pre .86
		Post	163	3.71	1.13		
	Control	Pre	164	3.55	1.08	Post .98	Post .88
		Post	164	3.55	1.14		
Ought-to L2 Self	Treatment	Pre	163	3.42	1.16	Pre .98	Pre .75
		Post	163	3.61	1.3		
	Control	Pre	164	3.36	1.13	Post .97	Post .85
		Post	164	3.43	1.34		
L2 Learning Experience	Treatment	Pre	163	4.29	0.96	Pre .98	Pre .88
		Post	163	4.47	1.09		
	Control	Pre	164	4.24	1	Post .97	Post .89
		Post	164	4.22	1.02		
Writing Self-Efficacy	Treatment	Pre	163	3.52	0.86	Pre .99	Pre .91
		Post	163	3.95	0.93		
	Control	Pre	164	3.63	0.81	Post .99	Post .92
		Post	164	4.10	0.84		

Table 2 Two-way ANOVA with interaction (Ideal L2 self).

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i> -ratio	<i>p</i> -value	Partial η^2
Treatment	1.02	1	1.02	0.48	.49	< .01
Time	1.12	1	1.12	3.96	<.05	.01
Treatment × Time	1.06	1	1.06	3.73	.05	.01

Table 3 Post analysis (Ideal L2 self).

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i> -ratio	<i>p</i> -value	Partial η^2
Treatment at Pre	< 0.01	1	< 0.01	< 0.01	.99	< .01
Treatment at Post	2.07	1	2.07	1.61	.21	< .01
Time in Control	< 0.01	1	< 0.01	< 0.01	.97	< .01
Time in Treatment	2.17	1	2.17	7.17	.01	.04

The interaction between treatment and time approached significance ($p = .05$) with a small effect size (Partial $\eta^2 = .01$). Despite the marginal significance of this interaction, post-hoc tests were conducted to explore potential differences further. This approach aimed to examine subtle interactions and offer a more thorough understanding of the data, as near-significant findings may reveal relevant trends and insights.

Table 3 provides the post-hoc analysis results, highlighting a significant interaction between the two factors. The analysis indicates that treatment alone had no significant effect at either the pre-test ($p = .99$) or post-test ($p = .21$) stages, with minimal effect sizes (Partial $\eta^2 < .01$). Time also showed no significant effect within the control group ($p = .97$), with no measurable effect size (Partial $\eta^2 < .01$). However, in the treatment group, time had a significant effect ($p = .01$) with a moderate effect size (Partial $\eta^2 = .04$), suggesting that ChatGPT usage produced a meaningful impact on students' Ideal L2 Self over time.

According to the results in Table 4, treatment alone did not have a significant effect ($p = .30$) and showed a negligible effect size (Partial $\eta^2 < .01$). Time, however, showed a significant effect ($p = .04$) with a small effect size (Partial $\eta^2 = .01$), indicating that changes over time were meaningful for both groups. The interaction between treatment and time was not significant ($p = .33$) with no notable effect size (Partial $\eta^2 < .01$), suggesting that ChatGPT use had no impact on students' Ought-to L2 Self.

The results of L2 Learning Experience in Table 5 indicate that treatment alone had no significant effect ($p = .15$) with a very small effect size (Partial $\eta^2 < .01$). Time showed a marginally significant effect ($p = .10$) with a similarly small effect size (Partial $\eta^2 < .01$), suggesting slight changes over time. The interaction between treatment and time, however, was significant ($p < .05$) with a small effect size (Partial $\eta^2 = .01$), pointing to a meaningful interaction effect. Consequently, a post-hoc test was conducted, with the results displayed in Table 6.

The post-hoc test results indicate that treatment had no significant effect at the pre-test stage ($p = .64$, Partial $\eta^2 < .01$). However, at the post-test stage, treatment showed a significant effect

Table 4 Two-way ANOVA with interaction (Ought-to L2 self).

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F-ratio</i>	<i>p-value</i>	Partial η^2
Treatment	2.65	1	2.65	1.09	.30	< .01
Time	2.74	1	2.74	4.46	.04	.01
Treatment $\times$ Time	0.57	1	0.57	0.94	.33	< .01

Table 5 Two-way ANOVA with interaction (L2 Learning Experience).

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F-ratio</i>	<i>p-value</i>	Partial η^2
Treatment	3.58	1	3.58	2.13	.15	< .01
Time	1.10	1	1.10	2.78	.10	< .01
Treatment $\times$ Time	1.55	1	1.55	3.91	<.05	.01

Table 6 Post analysis (L2 Learning Experience).

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F-ratio</i>	<i>p-value</i>	Partial η^2
Treatment at Pre	0.21	1	0.21	0.22	.64	< .01
Treatment at Post	4.92	1	4.92	4.41	.04	.01
Time in Control	0.02	1	0.02	0.04	.83	< .01
Time in Treatment	2.62	1	2.62	7.33	<.01	.04

($p = .04$, Partial $\eta^2 = .01$). Time had no significant effect in the control group ($p = .83$, Partial $\eta^2 < .01$) but showed a significant effect in the treatment group ($p < .01$, Partial $\eta^2 = .04$). These findings suggest that using ChatGPT positively impacted students' L2 Learning Experience over time.

Table 8 presents the results for Writing Self-Efficacy, showing that treatment alone had no significant effect ($p = .15$) with a negligible effect size (Partial $\eta^2 < .01$). The time factor showed a marginal effect ($p = .10$) with a very small effect size (Partial $\eta^2 < .01$). The interaction between treatment and time was significant ($p < .05$) with a small effect size (Partial $\eta^2 = .01$), indicating a meaningful interaction effect between the two factors, which warranted further examination through a post-hoc test.

The post-hoc test results in Table 9 indicate that treatment had no significant effect at the pre-test stage ($p = .64$) with a negligible effect size (Partial $\eta^2 < .01$). At the post-test stage, however, treatment showed a significant effect ($p = .04$) with a small effect size (Partial $\eta^2 = .01$). Time had no significant effect in the control group ($p = .83$) with a negligible effect size (Partial $\eta^2 < .01$), whereas time in the treatment group showed a significant effect ($p < .01$) with a moderate effect size (Partial $\eta^2 = .04$). These results suggest that ChatGPT use was effective in enhancing students' writing self-efficacy.

Table 7 Two-way ANOVA with interaction (Writing Self-Efficacy).

Source	SS	df	MS	F-ratio	p-value	Partial η^2
Treatment	3.58	1	3.58	2.13	.15	< .01
Time	1.10	1	1.10	2.78	.10	< .01
Treatment $\times$ Time	1.55	1	1.55	3.91	<.05	.01

Table 8 Post analysis (Writing Self-Efficacy).

Source	SS	df	MS	F-ratio	p-value	Partial η^2
Treatment at Pre	0.21	1	0.21	0.22	.64	< .01
Treatment at Post	4.92	1	4.92	4.41	.04	.01
Time in Control	0.02	1	0.02	0.04	.83	< .01
Time in Treatment	2.62	1	2.62	7.33	<.01	.04

Table 9 Correlation of Writing Self-Efficacy with motivational factors.

Group	WSE	IL2	OL2	L2LE
Control	Pre	.47	.23	.47
	Post	.44	.25	.40
Treatment	Pre	.48	.19	.53
	Post	.50	.35	.60

Note. All values are $p < .05$

Table 9 presents the Pearson correlation coefficients (r) for Writing Self-Efficacy (WSE) in relation to three motivational factors—Ideal L2 Self (IL2), Ought-to L2 Self (OL2), and L2 Learning Experience (L2LE)—for both the control and treatment groups, measured before and after the intervention. In the control group, WSE shows moderate correlations with IL2 ($r = .47$) and L2LE ($r = .47$) at the pre-test stage, which slightly decrease at the post-test stage ($r = .44$ for IL2 and $r = .40$ for L2LE), while the correlation with OL2 slightly increases from $r = .23$ to $r = .25$. In the treatment group, WSE shows moderate correlations with IL2 ($r = .48$) and L2LE ($r = .53$) at pre-test, which both increase at post-test. Notably, the correlation with OL2 rises from $r = .19$ to $r = .35$, and with L2LE from $r = .53$ to $r = .60$. These results suggest that the treatment positively strengthened the correlations between WSE and the motivational factors, whereas the control group exhibited either stable or slightly reduced correlations over time.

Discussion

In addressing RQ1, the treatment group reported a higher mean score, whereas the control group reported no change in mean score. Although the interaction of the two factors showed

marginal significance, the post-hoc test identified one significant finding. This suggests that instructor-led ChatGPT use in an EFL writing class positively influenced students' Ideal L2 Self motivation, as students in the treatment group reported higher scores. Hence the answer to this RQ is "Yes, students do report higher Ideal L2 Self motivation after instructor-led GenAI usage in the classroom".

The Ideal L2 Self concept serves as a motivational driver by highlighting the gap between one's current state and their desired future self (Dörnyei, 2009). In this study, the treatment group's higher mean score indicates that students view this innovative tool as a valuable resource for academic support. This positive reinforcement helps enhance their self-perception. The tool's ease of use allows students to integrate it seamlessly into their studies, increasing their motivation and aligning with the Ideal L2 Self (Taguchi et al., 2009; Ryan, 2009). Conversely, the control group, which did not use ChatGPT, showed no change in their Ideal L2 Self scores. This may be due to a lack of engagement, as they experienced consistent lecture styles throughout the semester, even with peer feedback. The absence of new and interactive elements in their learning experience could have reduced their sense of inspiration and self-improvement. This finding aligns with research on AI-mediated language instruction, which has been shown to improve L2 motivation (Wei, 2023), and it replicates the findings of Huang and Mizumoto (2024a), where the use of ChatGPT enhanced students' Ideal L2 Self motivation. This, in turn, supports the idea that AI can foster more engaging and personalized learning experiences, leading to increased motivation.

In addressing RQ2, Ought-to L2 Self, showed no significant interaction between the two factors. This indicates that ChatGPT use does not significantly impact Ought-to L2 Self. While both groups reported slightly higher motivation in this area after participating in peer feedback or instructor-led GenAI-supported classes, the difference was not statistically significant between the two groups. Therefore, the answer to this RQ is "No, students do not report higher Ought-to L2 Self motivation after instructor-led GenAI usage in the classroom."

The Ought-to L2 Self represents external pressure or obligations in language learning, often seen as a sense of duty (Dörnyei, 2009). Although both groups showed slight increases in scores, these changes were not significant. A possible explanation is that all students were enrolled to meet mandatory credit requirements for graduation. This constant obligation to fulfill credit requirements, regardless of tool use, aligns with findings on the Ought-to L2 Self being influenced by instrumental factors (Taguchi et al., 2009). Thus, external pressures remained constant and did not significantly affect language-learning motivation, which aligns with previous findings that ChatGPT use does not impact students' Ought-to L2 Self motivation.

For RQ3, L2 Learning Experience, the importance of interaction between the two factors is highlighted. While students in the treatment group reported higher mean scores, the control group showed a slight decline. This suggests that instructor-led AI use in EFL writing classes positively impacts students' L2 Learning Experience and enhances it. Similar to RQ1, the answer to this RQ is "Yes, students do report a higher L2 Learning Experience motivation after instructor-led GenAI usage in the classroom."

The L2 Learning Experience concept involves attitudes toward learning closely associated with the immediate learning environment (Dörnyei, 2009). In this study, the treatment group's mean score was slightly higher, while the control group exhibited a lower mean score. These results are consistent with expectations, as the introduction of the tool aligns with students' current learning context, adding a new dimension to their environment, which aligns with Dörnyei's (2009) construct. For the treatment group, the tool's integration enriched their learning environment, leading to an elevated mean score. Meanwhile, the control group experienced a steady lecture style without new elements, resulting in a lower mean score. This finding also

supports previous research, demonstrating that ChatGPT positively influenced students' motivation in their L2 Learning Experience. Furthermore, this result can be attributed to the ability of AI tools to personalize learning and enhance the learning environment. As noted by Leong et al. (2024) and Yamaoka (2024), AI tools can tailor learning and function as personalized teaching assistants, fostering greater motivation and a more positive learning experience.

In addressing RQ4, two significant interactions were found between the two factors. Although both the control and treatment groups showed higher mean scores in the post-measurement, these two significant results further validate the tool's effectiveness. Therefore, it can be inferred that using GenAI positively impacts students' writing self-efficacy. Like in RQ1 and RQ3, the answer to this question is "Yes, students report higher writing self-efficacy following instructor-led GenAI use in the classroom."

The effectiveness of the GenAI tool in this study demonstrated how it can create a supportive classroom environment. By providing immediate feedback, GenAI eliminates the usual time lag students experience when waiting for feedback from teachers. Given that self-efficacy can help reduce foreign language anxiety (Teng, 2024a), these findings reinforce the value of GenAI as a contextual support tool that fosters self-efficacy (Teng, 2024c). Additionally, the study's findings align with Teng and Yang's (2022) conclusions on the role of GenAI in enhancing both self-efficacy and motivation. Therefore, incorporating GenAI into the classroom holds potential for creating a more supportive learning context.

To address RQ5, the correlations of writing self-efficacy with motivational factors among students using GenAI showed increases across all three factors. In contrast, students using peer feedback showed either stable or decreased correlations. This suggests that using GenAI further strengthens the relationship between writing self-efficacy and motivation. Therefore, the answer to this question is: "Yes, the correlations of writing self-efficacy with motivational factors increase after using GenAI in the language classroom."

Self-efficacy not only refers to students' belief in their ability to achieve specific goals (Bandura, 1997), but it also affects their performance and outcomes (Bandura, 1997; Schunk, 1991). In this study, students demonstrated enhanced writing self-efficacy, illustrating the tool's effectiveness when implemented appropriately. Furthermore, self-efficacy and motivation are highly correlated (Teng & Wu, 2023), so an increase in one tends to result in an increase in the other. Referring to the first three research questions, all three motivational factors showed increased mean scores, which aligns with the rise in writing self-efficacy. Additionally, this study demonstrated that the correlations between writing self-efficacy and motivation strengthened after the use of GenAI, further underscoring the tool's positive impact. This finding reinforces the correlation between motivation and self-efficacy reported by Teng and Wu (2023). Although previous studies, such as those by Huang and Mizumoto (2024), did not explore this specific aspect, the results align with expectations given the relationship between these two variables. This evidence suggests that the careful introduction of GenAI in a classroom setting can enhance students' motivation and self-efficacy simultaneously.

This study, along with previous studies, provides valuable insights into the pedagogical applications of GenAI, particularly in EFL writing classrooms. The findings suggest that instructor-led integration of tools like ChatGPT can be instrumental in fostering students' Ideal L2 Self-motivation, their perceived L2 Learning Experience, and their overall writing self-efficacy. This implies that educators should carefully consider incorporating these AI tools into their teaching practices, focusing on enhancing the learning environment and providing robust academic support. However, it's crucial to note that while AI can boost intrinsic motivation, it doesn't significantly impact students' Ought-to L2 Self-motivation, which stems from external pressures and obligations. This emphasizes that AI tools are most effective when used to

foster a genuine interest in learning rather than to fulfill external requirements. Additionally, the study advocates for a structured approach to using ChatGPT, including crafted prompts, preliminary paper-based tasks, and documentation of changes, to prevent overreliance and plagiarism (Huang, 2023). This underscores the need for educators to guide students towards responsible and effective AI usage in language learning.

This study aimed to validate previous findings, and while it did confirm all of them, several limitations remain. First, the study's generalizability is limited because it was conducted only in Japan with primarily first- and second-year undergraduate students. Second, the survey relied heavily on self-reported data, which can introduce potential bias. Third, students' ChatGPT usage outside the classroom was neither monitored nor accounted for, potentially impacting results. Fourth, data were collected over a single semester, limiting insights into long-term effects. Fifth, the study focused solely on writing classes, so the findings may not apply to other skills, such as speaking. Lastly, students' actual writing proficiency was not measured.

To address these limitations, future research should expand sample pools to include a more diverse population across different cultural and educational backgrounds, improving generalizability. Including students from various years and disciplines could also provide a broader perspective on the tool's effectiveness. Adding student interviews may further strengthen self-report measures. Although it may be challenging, tracking students' ChatGPT usage outside the classroom could yield insights into independent usage patterns, helping to differentiate classroom-based effects from external influences. Conducting longitudinal studies across multiple semesters would allow researchers to examine the effects of sustained GenAI exposure on students' motivation, self-efficacy, and skill development over time. Extending the study to speaking classes and comparing effects across different language skills could offer insights into the transferability of these tools. Finally, evaluating students' writing samples for complexity, accuracy, and fluency would provide more objective measures of improvement. By addressing these limitations, future research can yield a more comprehensive understanding of GenAI's role in language learning and offer more robust evidence of its impact on motivation and self-efficacy.

Conclusion

This study presents several pedagogical implications for EFL instructors integrating GenAI tools, such as ChatGPT, into writing classrooms. First, fostering motivation and a positive learning experience is essential. To encourage intrinsic motivation, instructors should frame AI as a tool that empowers learning, design activities that promote active engagement and exploration, and highlight ChatGPT's potential to provide personalized feedback. Second, developing critical thinking skills alongside AI proficiency is crucial. Instructors should position AI as a collaborative tool, not a substitute for critical thought. This can be achieved by integrating paper-based writing and reflective activities, and by promoting responsible, ethical AI usage. Third, creating a supportive and inclusive learning environment is key. This involves offering structured guidance, facilitating peer collaboration, and addressing equity and access to ensure all students can engage with AI meaningfully. Finally, promoting ongoing professional development is essential. Instructors should be encouraged to stay informed about AI advancements and to experiment with various AI tools and instructional approaches, which will help them better support their students in a rapidly evolving tech landscape.

In conclusion, this study confirms the potential benefits of GenAI in EFL education, particularly in enhancing students' internal motivation and writing self-efficacy and strengthening their correlations. While AI tools can transform the language-learning experience by supporting

students' Ideal L2 Self and enhancing their learning context, their role in affecting obligation-driven motivation is limited. Future research should address the study's limitations by broadening the sample and extending research to other skill areas and learning contexts. These findings advocate for a balanced and carefully monitored integration of AI in classrooms, aimed at supporting authentic, sustained interest in language learning.

References

- Ahmetovic, E., Becirovic, S., Dubravac, V., & Brdarevic-Celjo, A. (2023). The interplay between corrective feedback, motivation and EFL achievement in middle and high school education. *Journal of Language and Education*, 9(1), 26–41. <https://doi.org/10.17323/jle.2023.12663>
- Aizawa, I., Rose, H., Thompson, G., & Curle, S. (2020). Beyond the threshold: Exploring English language proficiency, linguistic challenges, and academic language skills of Japanese students in an English medium instruction programme. *Language Teaching Research*, 27(4), 837–861. <https://doi.org/10.1177/1362168820965510>
- Al-Hoorie, A. H. (2018). The L2 motivational self system: A meta-analysis. *Studies in Second Language Learning and Teaching*, 8(4), 721–754. <https://doi.org/10.14746/ssllt.2018.8.4.2>
- Al-Hoorie, A. H., Hiver, P., & In'nami, Y. (2023). The validation crisis in the L2 motivational self system tradition. *Studies in Second Language Acquisition*, 46(2), 307–329. <https://doi.org/10.1017/s0272263123000487>
- Allen, T. J., & Mizumoto, A. (2024). ChatGPT over my friends: Japanese EFL learners' preferences for editing and proofreading strategies. *RELJ Journal*. <https://doi.org/10.1177/00336882241262533>
- Athanassopoulos, S., Manoli, P., Gouvi, M., Lavidas, K., & Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research*, 3(2), 818–824. <https://doi.org/10.25082/AMLER.2023.02.009>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W H Freeman.
- Barrett, A., & Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20, 1–24. <https://doi.org/10.1186/s41239-023-00427-0>
- Barrot, J. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Cai, Q., Lin, Y., & Yu, Z. (2023). Factors influencing learner attitudes towards ChatGPT-assisted language learning in higher education. *International Journal of Human-Computer Interaction*, 40(22), 7112–7126. <https://doi.org/10.1080/10447318.2023.2261725>
- Cen, Y., & Zheng, Y. (2023). The motivational aspect of feedback: A meta-analysis on the effect of different feedback practices on L2 learners' writing motivation. *Assessing Writing*, 59, 100802. <https://doi.org/10.1016/j.asw.2023.100802>
- Cotton, D., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Csizér, K., & Kormos, J. (2009). Learning experiences, selves and motivated learning behaviour: A comparative analysis of structural models for Hungarian secondary and university learners of English. In Z. Dörnyei & E. Ushioda (Eds), *Motivation, language identity and the L2 self* (pp. 98–117). Multilingual Matters.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum.
- Dörnyei, Z. (2020). *Innovations and challenges in language learning motivation*. Routledge.
- Fang, Y. (2023). Exploring the effects of teachers' feed up and feed forward on Chinese senior high students' English writing self-efficacy and writing motivation. *Journal of Education Humanities and Social Sciences*, 13, 30–36. <https://doi.org/10.54097/ehss.v13i.7851>
- Guo, K., & Wang, D. (2023). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 29, 8435–8463. <https://doi.org/10.1007/s10639-023-12146-0>

- Hao, Z., Fang, F., & Peng, J. (2024). The integration of AI technology and critical thinking in English major education in China: opportunities, challenges, and future prospects. *Digital Applied Linguistics*, 1, 2256. <https://doi.org/10.29140/dal.v1.2256>
- Henry, A., & Liu, M. (2024). Jingle-Jangle fallacies in L2 motivational self system research: A response to Al-Hoorie et al. (2024). *Applied Linguistics*, 45(4), 738–746. <https://doi.org/10.1093/applin/amae041>
- Huang, J. (2023). Engineering ChatGPT prompts for EFL writing classes. *International Journal of TESOL Studies*, 5(4), 73–79. <https://doi.org/10.58304/ijts.20230405>
- Huang, J., & Mizumoto, A. (2024a). The effects of generative AI usage in EFL classrooms on the L2 motivational self system. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-13071-6>
- Huang, J., & Mizumoto, A. (2024b). Examining the relationship between the L2 motivational self system and technology acceptance model post ChatGPT introduction and utilization. *Computers and Education: Artificial Intelligence*, 100302. <https://doi.org/10.1016/j.caeai.2024.100302>
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28, 15873–15892. <https://doi.org/10.1007/s10639-023-11834-1>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Lai, C. Y., Cheung, K. Y., & Chan, C. S. (2023). Exploring the role of intrinsic motivation in ChatGPT adoption to support active learning: An extension of the technology acceptance model. *Computers and Education Artificial Intelligence*, 5, 100178. <https://doi.org/10.1016/j.caeai.2023.100178>
- Leong, J., Pataranutaporn, P., Danry, V., Perteneder, F., Mao, Y., & Maes, P. (2024). Putting things into context: Generative AI-enabled context personalization for vocabulary learning improves learning motivation. *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1–15). Association for Computing Machinery (ACM).
- Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, 10, 1–17. <https://doi.org/10.1080/2331186X.2023.2236469>
- McDonald, N., Johri, A., Ali, A., & Hingle, A. (2024). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *arXiv*. <https://doi.org/10.48550/arxiv.2402.01659>
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 100050. <https://doi.org/10.1016/j.rmal.2023.100050>
- Mizumoto, A., Shintani, N., Sasaki, M., & Teng, M. F. (2024). Testing the viability of ChatGPT as a companion in L2 writing accuracy assessment. *Research Methods in Applied Linguistics*, 3(2), 100116. <https://doi.org/10.1016/j.rmal.2024.100116>
- Mohamed, A. M. (2024). Exploring the potential of an AI-based chatbot (ChatGPT) in enhancing English as a foreign language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies*, 29, 3195–3217. <https://doi.org/10.1007/s10639-023-11917-z>
- Pajares, F., & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology*, 24(4), 390–405. <https://doi.org/10.1006/ceps.1998.0995>
- Papi, M., & Teimouri, Y. (2024). Manufactured crisis: A response to Al-Hoorie et al. (2024). *Studies in Second Language Acquisition*. <https://doi.org/10.1017/s0272263124000494>
- Ryan, S. (2009). Self and identity in L2 motivation in Japan: The ideal L2 self and Japanese learners of English. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 120–143). Multilingual Matters. <https://doi.org/10.21832/9781847691293-007>

- Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26(3–4), 207–231. <https://doi.org/10.1080/00461520.1991.9653133>
- Stapleton, P., & Leung, B. K. K. (2019). Assessing the accuracy and teachers' impressions of Google Translate: A study of primary L2 writers in Hong Kong. *English for Specific Purposes*, 56, 18–34. <https://doi.org/10.1016/j.esp.2019.07.001>
- 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, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Taguchi, T., Magid, M., & Papi, M. (2009). The L2 motivational self system among Japanese, Chinese and Iranian learners of English: A comparative study. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 66–97). Multilingual Matters.
- Teng, M. F. (2023). Scientific writing, reviewing, and editing for open-access TESOL journals: The role of ChatGPT. *International Journal of TESOL Studies*, 5(1), 87–91. <https://doi.org/10.58304/ijts.20230107>
- Teng, M. F. (2024a). Do self-efficacy belief and emotional adjustment matter for social support and anxiety in online English learning in the digital era? *Digital Applied Linguistics*, 1, 2227. <https://doi.org/10.29140/dal.v1.2227>
- Teng, M. F. (2024b). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies*, 6(3), 36–57. <https://doi.org/10.58304/ijts.20240304>
- Teng, M. F. (2024c). “ChatGPT is the companion, not enemies”: EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education: Artificial Intelligence*, 7, 100270. <https://doi.org/10.1016/j.caeai.2024.100270>
- Teng, M. F., & Wu, J. G. (2023). An investigation of learners' perceived progress during online education: Do self-efficacy belief, language learning motivation, and metacognitive strategies matter? *The Asia-Pacific Education Researcher*, 33(2), 283–295. <https://doi.org/10.1007/s40299-023-00727-z>
- Teng, M. F., & Yang, Z. (2022). Metacognition, motivation, self-efficacy belief, and English learning achievement in online learning: Longitudinal mediation modeling approach. *Innovation in Language Learning and Teaching*, 17(4), 778–794. <https://doi.org/10.1080/17501229.2022.2144327>
- Tseng, W., & Warschauer, M. (2023). AI-writing tools in education: If you can't beat them, join them. *Journal of China Computer-Assisted Language Learning*, 3(2), 258–262. <https://doi.org/10.1515/jccall-2023-0008>
- Van Horn, K. (2024). ChatGPT in English language learning: Exploring perceptions and promoting autonomy in a university EFL context. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, 28(1). <https://doi.org/10.55593/ej.28109a8>
- Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1261955>
- Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and Perceptions. *Languages*, 8(3), 212. <https://doi.org/10.3390/languages8030212>
- Yamaoka, K. (2024). ChatGPT's motivational effects on Japanese university EFL learners: A qualitative analysis. *International Journal of TESOL Studies*, 6(3), 24–35. <https://doi.org/10.58304/ijts.20240303>
- Yang, S., Liu, Y., & Wu, T. (2024). ChatGPT, a new “Ghostwriter”: A teacher-and-students poetic autoethnography from an EMI academic writing class. *Digital Applied Linguistics*, 1, 2244. <https://doi.org/10.29140/dal.v1.2244>
- Zhang, J., Zhang, L. J., & Zhu, Y. (2023). Development and validation of a genre-based second language (L2) writing self-efficacy scale. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1181196>
- Zheng, Y., Wang, Y., Liu, K. S., & Jiang, M. Y. (2024). Examining the moderating effect of motivation on technology acceptance of generative AI for English as a foreign language learning. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12763-3>

Author Bio

Jerry Huang is an English language instructor at Kansai University and originally from Los Angeles. He holds a master's degree from Hyogo University of Teacher Education and is a Ph.D. candidate in foreign language education and research at Kansai University. Jerry's teaching expertise lies in English language courses, and his research interests include motivation, AI, ChatGPT, and Languages Other Than English (LOTE).

Atsushi Mizumoto holds a Ph.D. in Foreign Language Education and is a professor in the Faculty of Foreign Language Studies and the Graduate School of Foreign Language Education and Research at Kansai University, Japan. His current research interests include corpus use for pedagogical purposes, learning strategies, language testing, and research methodology.

Appendix

Ideal L2 Self

- I can imagine myself living abroad and having a discussion in English.
- Whenever I think of my future career, I imagine myself using English.
- I can imagine a situation where I am speaking English with foreigners.
- I imagine myself as someone who is able to speak English.
- The things I want to do in the future require me to use English.

Ought-to L2 Self

- I study English because close friends of mine think it is important.
- Learning English is necessary because people surrounding me expect me to do so.
- I have to study English, because, if I do not study it, I think my parents will be disappointed with me.
- My parents believe that I must study English to be an educated person.

L2 Learning Experience

- I like the atmosphere of my English classes.
- I always look forward to English classes.
- I find learning English really interesting.
- I really enjoy learning English.

Writing Self-Efficacy

- I can correctly spell all words in my paragraph.
- I can correctly punctuate my paragraph.
- I can correctly use all parts of speech in my paragraph.
- I can write simple sentences with good grammar.
- I can correctly use singulars and plurals, verb tenses, prefixes, and suffixes.
- I can write a strong paragraph that has a good topic sentence and main idea.
- I can structure paragraphs to support ideas in the topic sentences.
- I can end paragraph with proper conclusions.
- I can get ideas across in a clear manner by staying focused without getting off topic.