







The role of generative AI in mediating L2MSS and engagement with written feedback in EFL learning: A structural equation modeling approach

Jerry Huang¹ (i) and Atsushi Mizumoto² (i)

Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

Corresponding author: Jerry Huang; Email: jerryyhuang@cc.kyoto-su.ac.jp

Abstract

This empirical study explores three aspects of engagement (affective, behavioral, and cognitive) in language learning within an English as a Foreign Language context in Japan, examining their relationship with AI utilization. Previous research has demonstrated that motivation positively influences AI usage. This study expands on that by connecting motivation with engagement, where AI usage serves as an intermediary construct. A total of 174 students participated in the study. Throughout the semester, they were required to use Generative AI (GenAI) to receive feedback on their writing. To prevent overreliance or plagiarism, carefully crafted prompts were selected. Students were tasked with collaboratively constructing essays during the semester using GenAI. At the end of the semester, students completed a survey measuring their motivation and engagement. Structural Equation Modeling was employed to reaffirm the previous finding that motivation influences AI usage. The results showed that AI usage impacts all three aspects of engagement. Based on these findings, the study suggests the pedagogical feasibility of implementing GenAI in writing classes with proper teacher guidance. Rather than being a threat, the use of this technological tool complements the role of human teachers and supports learning engagement.

Keywords: the L2 motivational self system; ideal L2 self; ought-to L2 self; L2 learning experience; generative AI; ChatGPT; technology acceptance model; affective engagement; behavioral engagement; cognitive engagement

Introduction

With the increasing integration of artificial intelligence (AI) in education, second language (L2) learning is undergoing a significant transformation. Generative AI

© The Author(s), 2025. Published by Cambridge University Press.

 $^{^1}$ Faculty of Cultural Studies, Kyoto Sangyo University Motoyama, Kamigamo, Kita-ku, Kyoto, Japan and

²Faculty of Foreign Language Studies, Kansai University, Suita-shi, Osaka, Japan

(GenAI), in particular, is poised to revolutionize the learner experience, building upon existing technologies like machine translation that are already widely used. Unlike earlier technologies that primarily focused on translation or rote memorization, GenAI offers a more dynamic and interactive learning environment. Its ability to generate human-like text, provide personalized feedback, and create immersive learning scenarios has the potential to significantly enhance language acquisition. This shift in language learning calls for a deeper understanding of student engagement and its role in the learning process. Previous research by Huang and Mizumoto (2024a, 2024b, 2024c) demonstrated that using GenAI with appropriate measures not only supports and maintains learners' motivation but also improves writing self-efficacy. Additionally, several studies have shown that technology use positively influences engagement (e.g., Chen et al., 2009; Katyara et al., 2023; Maričić & Lavicza, 2024; Schindler et al., 2017), which is a key driver of language learning (Iwaniec & Khaled, 2024). This underscores the importance of investigating GenAI's role in the language learning classroom further. While previous research has explored the impact of technology on engagement and the relationship between motivation and AI use, limited studies have examined the connection between motivation, engagement, and AI utilization in the context of L2 learning. Given that engagement is closely tied to motivation (Iwaniec & Khaled, 2024), this empirical study aimed to explore the intricate relationship between motivation, engagement, and AI utilization — with a specific focus on AI-generated written feedback (WF) - in an English as a Foreign Language (EFL) context in Japan.

Literature review

L2 motivation self system

L2 Motivational Self System (L2MSS) was introduced by Dörnyei in 2009 as a comprehensive framework to better understand language learning motivation (Boo et al., 2015), expanding on previous theories like Gardner's (1985) integrative/instrumental dichotomy. This model is particularly relevant in contexts like EFL, where students often lack meaningful contact with native speakers. The L2MSS is divided into three distinct constructs, namely the Ideal L2 Self (IL2), the Ought-to L2 Self (OL2), and the L2 Learning Experience (L2LE). These constructs allow for deeper insight into how learners' future-oriented goals and direct learning experiences affect their motivation. At the heart of L2MSS lies the IL2, a concept derived from the idea of possible selves (Markus & Nurius, 1986). This construct refers to a learner's vision of the person they aspire to become by mastering the language, often tied to professional and personal aspirations. Studies have demonstrated that a strong IL2 often correlates with higher motivation and improved language proficiency (e.g., Al-Hoorie, 2018; Dunn & Iwaniec, 2021). Promotion-focused goals, where learners are driven by positive outcomes, align with this concept, suggesting that students with a well-developed IL2 tend to put in more effort to achieve their goals. However, some recent studies suggest that this correlation may not be universal (Takahashi & Im, 2020), indicating the need for further research to fully understand this construct's impact on language learning. The OL2, on the other hand, represents a learner's sense of obligation to learn

the language to meet external expectations, such as avoiding failure or fulfilling societal responsibilities (Dörnyei & Taguchi, 2009). In contrast to the IL2, this construct is more prevention-focused, where learners are motivated by a desire to avoid negative consequences. While traditionally seen as less impactful than the IL2, studies in Asian contexts, such as Japan (Suzuki, 2014), suggest that the OL2 can still hold significant motivational power, particularly in educational settings where English is a compulsory subject. Researchers continue to debate whether this construct should be further divided into internally and externally motivated components. The L2LE shifts the focus from future-oriented goals to the immediate experiences learners have while engaging with the language. This construct encompasses various factors, including interactions with teachers and peers, classroom environment, and study methods. Although the L2LE is recognized as an important motivational factor (e.g., Lamb, 2012; Takahashi & Im, 2020; Teimouri, 2016), it has received less attention compared to the other constructs. A recent study by Huang et al. (2024) demonstrated the potent factor of L2LE within the L2MSS and its influence on IL2 and OL2. They concluded that L2LE, encompassing the immediate learning environment, should serve as the strongest predictor for intended effort and proficiency. Research indicates that the quality of students' learning experiences plays a critical role in their motivation to continue learning. Despite its potential importance, the L2LE remains underexplored, and its broad scope has prompted calls for more detailed studies to break down its components and better understand its influence on language learning. Researchers investigating the L2MSS often focus on the IL2 and OL2 and their connections to intended effort, as seen in studies like Yashima et al. (2017) and Papi (2010). While these constructs are strong predictors of intended effort, their influence on actual achievement is less certain. Fewer studies, however, have examined the role of the L2LE in relation to these constructs. Some research suggests a strong relationship between the L2LE and both the IL2 and OL2, while others, such as Csizér and Kormos (2009), report weaker associations. Yashima et al. (2017) explored the impact of L2LE, redefined as communication orientation and grammar-translation orientation, and found positive influences on both IL2 and OL2, highlighting the importance of the L2LE, particularly in the Japanese context.

Engagement

In second language acquisition (SLA) research, learning engagement is typically categorized into three main dimensions: affective engagement (AE), behavioral engagement (BE), and cognitive engagement (CE). AE relates to students' emotional reactions and attitudes toward learning and feedback. In the context of WF, researchers found that students generally react well to feedback that is detailed and balanced, providing motivation without overwhelming corrections, which boosts their motivation and self-efficacy (Ene & Yao, 2021; Purnomo & Pahlevi, 2021). However, the type of feedback affects engagement, with feedback focused on form often receiving more positive responses than content-focused feedback (Fan & Xu, 2020). The tone and delivery of feedback are important in influencing emotional responses and maintaining engagement (Yu et al., 2020). Additionally, language proficiency plays a role in

how students engage, as high- and low-proficiency learners show different reactions, particularly to localized feedback (Cheng & Liu, 2022). In EFL settings, students may respond better to feedback from native English-speaking teachers compared to nonnative teachers (Ene & Yao, 2021). AE is closely connected to both CE and BE, forming a complex, dynamic relationship (Cheng & Liu, 2022; Fan & Xu, 2020). BE is the observable actions and efforts that learners take in response to WF, commonly measured by the extent and quality of their revisions (Fu et al., 2024). Although revisions are a primary indicator, CE and AE may still occur even if feedback does not lead to immediate changes (Fu et al., 2024). Proficiency also impacts how students engage, with higher proficiency learners showing greater cognitive involvement in response to teacher feedback (Cheng & Liu, 2022). In peer feedback situations, BE tends to be most evident, followed by CE and AE (Farsani & Aghamohammadi, 2021). Activities that require deeper cognitive processing generally lead to a higher uptake of written corrective feedback (WCF) during revisions, although results can differ depending on the type of error (Park & Ahn, 2022). Factors such as the quality of feedback, proficiency level, and well-structured peer review activities also significantly affect engagement (Fu et al., 2024). These insights point to the intricate and interconnected nature of BE, CE, and AE in L2 acquisition and EFL contexts. CE refers to the mental processes learners utilize when engaging with WF. This includes understanding, analyzing, and applying feedback, which encourages reflection on language use and fosters deeper cognitive processing, potentially leading to better writing outcomes. Activities that stimulate deeper CE are associated with a higher uptake of WCF during revisions (Park & Ahn, 2022). Research shows that CE is most common with teacher feedback, while BE is more typical in peer feedback scenarios (Farsani & Aghamohammadi, 2021). Language proficiency also shapes how learners engage, with differences in how lowand high-proficiency learners handle localized feedback, though both groups engage similarly with global feedback (Cheng & Liu, 2022). Additionally, self-regulation abilities influence engagement, with stronger self-regulators exhibiting higher levels of CE, BE, and AE compared to their less-skilled peers (Yang & Zhang, 2023). These findings emphasize the complexity of student engagement with WCF and the need to tailor feedback to maximize cognitive involvement and support writing improvement. Research has shown that the three factors of engagement — affective, behavioral, and cognitive — are closely linked and inseparable. AE leads to increased BE and CE, as demonstrated by Ebadi et al. (2024), indicating that when learners experience positive emotions, they are more likely to actively participate and invest mental effort in language learning. BE is also closely tied to AE and CE aspects, as shown in Cheng and Zhang's (2024b) study, which found that students who actively participated were more likely to develop positive attitudes and engage in deeper cognitive processing of language input. CE, in turn, influences both AE and BE. Pearson's (2024) systematic review concluded that learners who are cognitively engaged tend to experience more positive emotions and show greater participation. In summary, these three engagement factors are intertwined — enhancing one boosts the others. Therefore, it is essential for language teachers to create environments that foster all three aspects of engagement.

Motivation and engagement

The construct of motivation and engagement are closely intertwined, forming a dynamic and reciprocal relationship. Motivation is often considered a broad concept encompassing interest and engagement (Renninger et al., 2018). Engagement, in contrast, has been described as "motivation-plus," acting as an extension of motivation and its critical application (Mercer & Dörnyei, 2020). Research underscores their interplay: motivation drives higher engagement, which in turn enhances motivation, suggesting a cyclical relationship (Martin et al., 2017; Sulis, 2020). This reciprocal relationship has been observed even before the integration of GenAI into education. In the context of language learning, motivated students tend to show higher levels of engagement, and engaged students, in turn, experience greater motivation (Abdollahzadeh et al., 2022). Studies have consistently found that motivation fosters greater engagement and improved performance in foreign language learning (Kanellopoulou & Giannakoulopoulos, 2020; Noels et al., 2019). Additionally, engagement, particularly CE, has been identified as a predictor of motivation among EFL learners (Ghelichli et al., 2020). A key question in understanding motivation and engagement is their sequence: do they influence each other, and if so, which comes first? Some researchers argue that motivation precedes engagement. Reeve (2012) posits that students' internal motivational resources enhance classroom engagement. Similarly, Reeve et al. (2004) found that teachers trained in autonomy-supportive practices exhibited behaviors that increased students' engagement. Anderman and Patrick (2012) propose that motivation, through goals, precedes various types of engagement, such as cognitive (e.g., self-regulation), emotional (e.g., positive feelings about school), and BE (e.g., effort). Other researchers make distinctions between will (motivation) and skill (engagement) (Cleary & Zimmerman, 2012; Covington, 2000). Schunk and Mullen (2012) describe engagement as the manifestation of motivation, while Voelkl (2012) sees engagement as serving as an intermediary between motivation and achievement. Pekrun and Linnenbrink-Garcia (2012) similarly suggest that engagement mediates the relationship between emotion and achievement, and Ainley (2012) argues that motivation (via interest) leads to achievement through engagement. Previous empirical studies further clarify this relationship. Froiland and Worrell (2016) demonstrated that intrinsic motivation positively influences academic performance indirectly through classroom engagement. Their findings were replicated among African American and Latino students. In a longitudinal study, Froiland and Davison (2016) showed that intrinsic motivation in mathematics predicted engagement (e.g., enrollment in advanced mathematics courses), which subsequently led to higher achievement. These findings collectively suggest that motivation serves as the driving force behind engagement, though Reschly and Christenson (2012, p. 14) caution that "motivation is necessary but not sufficient for engagement." Hence, in this study, we acknowledge the intricate relationship between motivation and engagement, where some interpret their relationship as cyclical. However, we emphasize the predictive power of motivation towards engagement, particularly with the newly introduced GenAI, which can lead to higher engagement as an intermediary, but is not necessarily required for engagement to promote motivation. Motivation influences engagement in several ways. For example, L2LE, the strongest predictor in the L2MSS, encompasses the immediate

learning environment. One study in high school by Shernoff et al. (2014), using the experience sampling method, demonstrated that the learning environment can predict engagement, although this study focused on five different subject areas. Similarly, research by Saeed and Zyngier (2012) on grade five and six students and Liu et al. (2024) on sophomore university students both provided solid evidence of motivation influencing engagement. Therefore, fostering students' motivation is crucial in promoting sustained engagement and successful language acquisition (Noels et al., 2019). As shown by Huang and Mizumoto (2024b), the use of GenAI can help sustain and enhance students' motivation. Consequently, even with the introduction of GenAI in the classroom, students' engagement is likely to increase as well.

Motivation and GenAI utilization

Although the use of GenAI in EFL classrooms is still in its early stages, several studies have already shown a positive correlation between AI and student motivation. One study demonstrated that AI-mediated context positively influenced students' motivation, though its effect on learning performance was inconclusive (Leong et al., 2024). Similarly, another study found that AI-mediated instruction not only improved English learning motivation but also enhanced self-regulated learning and learning outcomes (Yang, 2024). Both studies focused on how AI affects students' motivation, rather than the other way around. Zheng et al. (2024), using the Unified Theory of Acceptance and Use of Technology 2, found that motivation is a strong predictor of behavioral intention, which ultimately leads to actual technology usage. Lai et al. (2023) also observed that intrinsic motivation influences behavioral intention in the Technology Acceptance Model after the classroom adoption of ChatGPT. Similarly, Huang and Mizumoto (2024a) concluded that OL2 motivation significantly affects AI usage in language learning, although they emphasized the need for further investigation into other constructs within the L2MSS that may have a similar impact.

GenAI and engagement in language learning

ChatGPT has shown promise in educational settings, benefiting both students and teachers. A recent systematic review by Lo et al. (2024) on student engagement with ChatGPT found mixed results regarding AE, with positive aspects like satisfaction and interest, alongside negative ones like disappointment and anxiety. There is limited yet significant evidence of BE, including instances of both engagement and disengagement, with some use cases involving academic dishonesty. CE was noted, but it was relatively weak, with improvements in understanding but a decline in critical thinking. The review covered not only language learning but also other educational areas. An intervention study by Wang and Xue (2024) focusing on Chinese EFL students using AI-powered chatbots reported improvements in all three areas of engagement. The AI chatbots played a key role in enhancing these engagements. Another study by Rad et al. (2023) explored the use of an AI tool in L2 writing classes and found that students significantly improved their writing outcomes, engagement, and feedback literacy. Two recent studies (Huang & Teng, 2025; Teng & Huang, 2025) examined the effect of

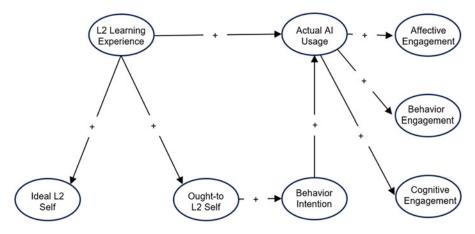


Figure 1. Hypothetical pathways.

ChatGPT utilization on engagement in China and Japan, concluding that all three factors increased following its use. These findings suggest the need for further research to explore how AI tools like ChatGPT can effectively support student engagement in language learning.

Research question

Previous research indicates that motivation influences the use of technology, which in turn affects student engagement. Applying the principle of mediation, we can infer that motivation impacts the use of AI technology and indirectly affects student engagement through its usage. We also acknowledge the versatility of GenAI, which includes capabilities for generating images, videos, and music. However, this study only utilized the text-based functions of ChatGPT; therefore, the actual AI usage in this study is defined as the use of the ChatGPT chatbot. Figure 1 summarizes these previous findings and illustrates the hypothetical pathways for this study. This empirical study aimed to validate these findings by addressing the following research question:

RQ: Does GenAI usage mediate the relationship between motivation and engagement?

Material and methods

Participants

The study involved 174 second-year students from a private university in the Kansai region of Japan, with 79 males and 95 females. All participants were enrolled in a required advanced reading and writing course, selected using convenience sampling. Despite their English proficiency not being formally evaluated, it was assumed to range from A2 to B2 on the Common European Framework of Reference (CEFR), based on previous research by Aizawa et al. (2020). Additionally, all participants had completed at least eight years of compulsory English education in primary and secondary school,

as well as one year of mandatory English instruction during their first year of university. While all students completed the course and required assignments, including preand post-writing, six students did not complete the survey. Consequently, only 168 students were included in the Structure Equation Modeling (SEM) analysis.

Instruments

In this study, a set of 37 questions, accessible on OSF (https://osf.io/3rs8t/), was adapted from the work of Taguchi et al. (2009), Liu and Ma (2023), and Cheng and Zhang (2024a). Taguchi et al's (2009) L2MSS questionnaire, originally developed for Japanese participants, is fully available in Japanese in Dörnyei (2003). These questions evaluated three core aspects of the L2MSS: IL2 (5 items), OL2 (4 items), and L2LE (4 items). Liu and Ma's (2023) survey was designed to measure two constructs of the Technology Acceptance Model (TAM): Behavioral Intention (3 items) and Actual Usage (6 items). For clarity, Actual Usage refers to the Actual AI Usage, which will be used hereafter. Cheng and Zhang's (2024a) instrument was modified to assess three dimensions of engagement after using GenAI: affective (4 items), behavioral (6 items), and cognitive (6 items). The questionnaire was developed in Japanese to guarantee precise linguistic comprehension and ensure cultural relevance. A pilot test was conducted with a small group of participants to confirm the clarity and appropriateness of the questions. Based on their feedback, necessary revisions were made to refine the questionnaire. Additionally, back-translation methods were employed to verify the accuracy of the Japanese content against the original English version. These steps ensured that the questionnaire accurately captured the intended constructs without linguistic ambiguity. To ensure a diverse order and uphold the integrity of the survey, the university's online learning management system randomized the sequence of the questions so that each participant received them in a different order. However, all participants completed the full set of 37 questions. If a survey was incomplete, a prompt was triggered, reminding participants to finish it. The survey was administered in Japanese, and responses were recorded using a 6-point Likert scale, where 1 represented "strongly disagree" and 6 represented "strongly agree."

Writing instruction

Following the same collection method as Huang and Mizumoto (2024a, 2024b, 2024c), students used GenAI, specifically ChatGPT, during two writing workshops throughout the semester, each lasting two weeks. These workshops were integrated into the regular academic semester and were taught by three different teachers, encompassing a total of 10 distinct classes. All participating teachers underwent standardized training on the integration of GenAI in writing workshops. This training included structured modules covering the functionality of GenAI tools, guidelines for workshop implementation, and best practices for facilitating discussions on AI-assisted writing. To ensure consistency, pre-training and post-training assessments were conducted, confirming the uniformity of instructional delivery and minimizing variations across teachers. In groups of five, students collaboratively crafted essays consisting of five paragraphs: an introduction, three body paragraphs, and a conclusion. Each student handled a specific

paragraph, with essay topics aligned with class readings. Week one of each workshop focused on creating flow charts, outlines, and drafts. Unfinished sections became homework assignments. In week two, students brought their typed paragraphs to class for ChatGPT feedback, guided by carefully crafted prompts (see Huang, 2023). These three distinct prompts were specifically designed to review different sections of an essay — introduction, body paragraphs, and conclusion — against detailed criteria. The criteria for the introduction include the hook, background information, thesis statement, organization, and clarity. For the body paragraphs, the criteria are topic sentence, development/support, organization/structure, transitions, analysis/critical thinking, and clarity/coherence. For the conclusion, the criteria encompass the restatement of the thesis, summary of key points, closing thought, connection to the introduction, and concluding sentence. The prompts provided feedback for improvement without rewriting the content. To avoid over-reliance on GenAI and ensure accuracy, students were required to submit the feedback they received from ChatGPT via university's online learning management system for evaluation by the teachers. Additionally, students had to document any changes on paper to demonstrate their work and prevent copy-pasting. Final group essays and revised individual drafts were submitted at the end of week two.

Data analysis

In our study's data analysis segment, we utilized R software (version 4.3.3) for a careful examination of the collected data. We commenced with basic descriptive statistics to capture the dataset's central tendencies and variability comprehensively. Initially, we tested the assumption of multivariate normality using the Henze–Zirkler test, which was not satisfied. Consequently, we adopted the weighted least squares means and variance adjusted estimation for robust parameter estimates (Beauducel & Herzberg, 2006). Further, we evaluated model fit indices to ensure our statistical models align with our theoretical frameworks, using SEM. This step was crucial to affirming the adequacy of our models in explaining the observed data. Next, we looked into the relationships between constructs to provide a nuanced analysis of the dataset's structural relationships.

For reproducibility and transparency, we have made the data and R code accessible on OSF (https://osf.io/3rs8t/).

Results

Descriptive statistics

To illustrate data trends, we examined the distribution of all eight constructs. As shown in Table 1, the mean scores for the constructs ranged from 3.50 to 4.48, with standard deviations between 0.85 and 1.37. Skewness values ranged from -0.55 to -0.05, and kurtosis values ranged from -0.96 to 0.53. Based on Kline's (2016) recommended cut-off values of ± 3.0 for skewness and ± 8.0 for kurtosis, the constructs displayed characteristics of a normal distribution. Additionally, Cronbach's alpha values ranged from 0.77 to 0.91, indicating a high level of internal consistency and reliability of the measurements. Although the Cronbach's alpha for Behavioral Intention and BE were

Table 1. Descriptive statistics

Construct	Mean	SD	Skewness	Kurtosis	α
IL2	3.70	1.12	-0.05	-0.67	.87
OL2	3.60	1.30	-0.15	-0.80	.84
L2LE	4.47	1.09	-0.55	-0.12	.90
BI	4.27	1.07	-0.53	-0.16	.77
AU	3.50	1.37	-0.05	-0.96	.91
AE	4.35	1.03	-0.37	-0.35	.85
BE	4.29	0.85	-0.46	0.36	.77
CE	4.48	0.85	-0.45	0.53	.84

Note: The abbreviations used in the data table are as follows: IL2 (Ideal L2 Self), OL2 (Ought-to L2 Self), L2LE (L2 Learning Experience), BI (Behavioral Intention), AU (Actual AI Usage), AE (Affective Engagement), BE (Behavioral Engagement), and CE (Cognitive Engagement)

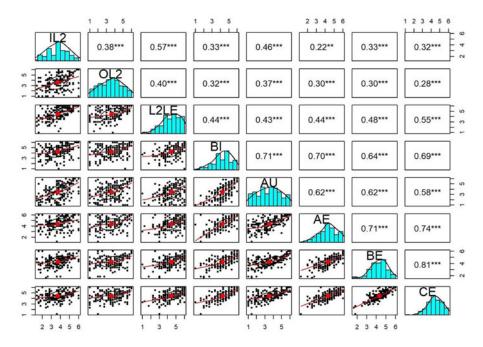


Figure 2. Histograms and correlation coefficients. Note: N=168.****p < .001; ***p < .01

the lowest at 0.77, it still demonstrated internal consistency, as argued by Dörnyei and Taguchi (2009, p. 95).

Correlation coefficients

Figure 2 showcases the Pearson correlation coefficients (r) among the variables. The observed correlation coefficients display considerable strength, indicating a strong

Table 2. Goodness-of-fit indices

Hypothesized model	RMSEA	SRMR	CFI	TLI
	0.048	0.069	0.981	0.979

Note: RMSEA = root mean squared error of approximation; SRMR = standardized root mean square residual; CFI = comparative normed fit index; TLI = Tucker-Lewis index.

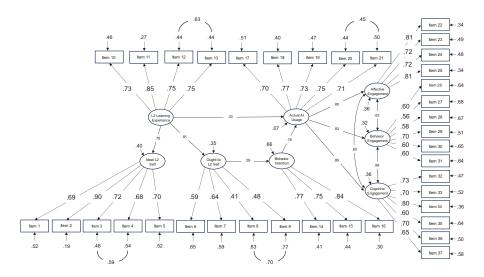


Figure 3. Path diagram of structure equation modeling results. *Note:* N=168. Standardized coefficients are shown for all paths (p<.05).

level of interrelatedness within the dataset. This finding suggests the data is highly suitable for multivariate analyses, such as SEM.

Goodness-of-fit indices

The goodness-of-fit indices shown in Table 2 demonstrate a robust model fit. Given the data did not meet the multivariate normality assumption, we used a robust estimation method. Consequently, we report robust fit indices, detailed in the supplementary material on OSF (https://osf.io/3rs8t/). The RMSEA and SRMR values of 0.048 and 0.069, respectively, are well below the recommended thresholds of RMSEA < 0.07 and SRMR < 0.08, indicating strong agreement between the model's predictions and the observed data (Hooper et al., 2008). Similarly, the CFI and TLI values of 0.981 and 0.979, respectively, approaching 1, suggest a good fit (Hooper et al., 2008). Overall, these goodness-of-fit indices support a solid model fit, aligning with our research objectives and providing a reliable basis for our conclusions.

SEM

The SEM diagram (Figure 3) successfully verified the influence of motivation on GenAI usage in the classroom, with usage further impacting students' engagement. The L2LE

has a moderate effect on Actual AI Usage ($\beta=.33$). Additionally, L2LE indirectly affects Actual AI Usage through OL2 and Behavioral Intention. The diagram also shows that Actual AI Usage strongly influences all three types of engagement — affective, behavioral, and cognitive — with respective correlations ($\beta=.80,\,\beta=.83,$ and $\beta=.80$). To improve the model fit, error variances were adjusted due to item similarities. Items 3 and 4 both address the speaker's ability to speak English, while Items 8 and 9 reflect the speaker's motivation to study English driven by parental expectations. Items 12 and 13 convey a positive attitude towards learning English, and Items 20 and 21 emphasize the use of ChatGPT to improve English language skills. Furthermore, the three engagement factors showed covariance (.63, .99, and .65), underscoring the interrelatedness of these aspects of engagement.

Discussion

Based on the SEM diagram, it can be said that GenAI mediates the relationship between motivation and engagement both directly and indirectly. The motivational construct L2LE directly influences Actual AI Usage ($\beta = .33$), aligning with previous research indicating a positive correlation between GenAI usage and motivation (e.g., Leong et al., 2024; Yang, 2024). Indirectly, motivation affects GenAI through the OL2 construct and Behavioral Intention. The pathway from L2LE to OL2 is strong ($\beta = .81$), consistent with Yashima et al.'s (2017) findings. OL2 then influences the Behavioral Intention construct ($\beta = .59$), further supported by previous research (e.g., Lai et al., 2023; Zheng et al., 2024). Although this finding deviates from Huang and Mizumoto's (2024a) previous results, which noted OL2's influence on Actual AI Usage, this empirical study overcomes the limitation of a small sample size. Furthermore, Behavioral Intention influences Actual AI Usage ($\beta = .59$), following the same pathway as the original technology acceptance model, and verified in Huang and Mizumoto's (2024a) study. Therefore, without a doubt, motivation influences the Actual AI Usage. The Actual AI Usage influences all three factors of engagement. Based on the SEM diagram, the Actual AI Usage demonstrated strong influence on AE ($\beta = .80$), BE ($\beta = .83$), and CE (β = .80). These findings align with previous research (e.g., Lo et al., 2024; Rad et al., 2023; Wang & Xue, 2024), where AI utilization in classrooms positively affects all three factors of engagement. Furthermore, the SEM pathways illustrate that all three factors are interrelated, with AE showing a covariance of .63 with BE and .65 with CE. Additionally, BE demonstrated a covariance of .99 with CE. These findings are also consistent with previous research (e.g., Cheng & Zhang, 2024b; Ebadi et al., 2024; Pearson, 2024), which suggests that the three constructs are inseparable; an increase in one result in an increase in the others. It is also worth mentioning that even after the introduction of GenAI usage in the classroom, the intercorrelation among the three factors of engagement remains intact. Thus, it is generally safe to say that the Actual AI Usage in the classroom mediates students' motivation and engagement, based on the empirical evidence of this study. While L2LE is considered the most crucial motivational factor of the L2MSS (e.g., Lamb, 2012; Takahashi & Im, 2020; Teimouri, 2016) and includes various elements like classroom interactions, environment, and learning methodology, the introduction of GenAI fits perfectly within this construct. Not only students but also teachers interact with GenAI in the classroom environment, necessitating a

new teaching pedagogy that incorporates this technology. Despite concerns from the research community about overreliance and plagiarism, these issues can be mitigated if educators are well trained and informed. This empirical study demonstrates overcoming these obstacles through carefully crafted prompts. Falling back to traditional paper and pencil methods to avoid GenAI would be akin to an ostrich burying its head in the sand. Rather than viewing GenAI as a can of worms, educators should leverage it to enrich students' L2LE. Instead of seeing GenAI as an adversary, it should be embraced as a companion (Teng, 2024). Writing in SLA poses significant challenges for EFL learners, particularly when it comes to receiving feedback from teachers, which can also place an extra burden on educators. Feedback is essential for students to improve their writing skills, yet the traditional approach can be taxing for both parties. This study demonstrated how GenAI can assist teachers by reducing their workload, while ensuring that every student still receives effective feedback. Students might feel more comfortable receiving feedback at their own pace, which can reduce anxiety. The results showed an improvement in writing coherence. This finding highlights the effectiveness of GenAI in enhancing writing coherence, creating a win-win situation for both students and teachers. By implementing GenAI more widely, students can receive valuable feedback without placing additional strain on teachers. Furthermore, teachers can craft prompts that focus on other aspects of writing, such as accuracy, complexity or fluency, enhancing the learning experience. In essence, integrating GenAI into the feedback process can significantly benefit EFL learners and educators alike. Despite these positive outcomes, this study has limitations. First, convenience sampling was used, limiting the generalizability of the findings to other populations. The results may not apply to students of different ages, locations, or institutions. Second, this study measured students' writing performance solely in terms of cohesion; other important aspects, such as accuracy, complexity, and fluency, were not considered. Third, the study was conducted over a single semester, a relatively short period. It is possible that the observed effects of GenAI on motivation, engagement, and writing performance might not persist over a longer timeframe. Finally, the study focused specifically on ChatGPT, a single GenAI tool, so the findings may not extend to other GenAI tools, such as Google's Gemini or Microsoft's Copilot. To address these limitations, future studies should aim for more diverse and representative sampling across different age groups, geographic locations, and educational institutions to enhance generalizability. Expanding the scope of writing performance measures to include aspects like accuracy, complexity, and fluency would provide a fuller picture of students' development. Conducting longitudinal research over multiple semesters could reveal whether the effects of GenAI on motivation, engagement, and performance are sustained over time. Additionally, including multiple GenAI tools in the study, such as Google's Gemini and Microsoft's Copilot, would allow for comparative analysis, offering insights into the broader applicability of the findings across various platforms.

Conclusion

This empirical study confirmed that GenAI, like ChatGPT, can positively influence motivation and engagement in EFL learning when thoughtfully integrated into language classrooms. Findings indicate that AI usage supports emotional, behavioral,

and cognitive engagement, reinforcing the connection between motivation and active participation. The results highlight the potential of AI to create a more interactive and stimulating learning environment, encouraging students to become more involved in their studies. Future studies could expand on this research by examining diverse AI tools and additional writing metrics over longer periods, providing deeper insight into AI's educational role. Pedagogical implications of integrating GenAI include a need for teacher training, especially in designing effective prompts to maximize AI's benefits (Kohnke et al., 2023). By doing so, educators can harness GenAI's advantages, enhancing learning experiences for EFL students and creating more engaging classrooms.

References

- Abdollahzadeh, E., Farsani, M. A., & Zandi, M. (2022). The relationship between L2 motivation and transformative engagement in academic reading among EAP learners: Implications for reading self-regulation. *Frontiers in Psychology*, 13, 944650. doi:10.3389/fpsyg.2022.944650
- Ainley, M. (2012). Students' interest and engagement in classroom activities. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), Handbook of research on student engagement (pp. 283–302). Springer.
- 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. doi: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. doi:10.14746/ssllt.2018.8.4.2
- Anderman, E. M., & Patrick, H. (2012). Achievement goal theory, conceptualization of ability/intelligence, and classroom climate. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), Handbook of research on student engagement (pp. 173–191). Springer.
- Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. Structural Equation Modeling: A Multidisciplinary Journal, 13(2), 186–203. doi:10.1207/s15328007sem1302_2
- Boo, Z., Dörnyei, Z., & Ryan, S. (2015). L2 motivation research 2005–2014: Understanding a publication surge and a changing landscape. *System*, 55, 145–157. doi:10.1016/j.system.2015.10.006
- Chen, P. D., Lambert, A. D., & Guidry, K. R. (2009). Engaging online learners: The impact of Web-based learning technology on college student engagement. *Computers and Education*, 54(4), 1222–1232. doi:10. 1016/j.compedu.2009.11.008
- Cheng, X., & Liu, Y. (2022). Student engagement with teacher written feedback: Insights from low-proficiency and high-proficiency L2 learners. System, 109, 102880. doi:10.1016/j.system.2022.102880
- Cheng, X., & Zhang, L. J. (2024a). Engaging secondary school students with peer feedback in L2 writing classrooms: A mixed-methods study. *Studies in Educational Evaluation*, 81, 101337. doi:10.1016/j.stueduc. 2024.101337
- Cheng, X., & Zhang, L. J. (2024b). Examining second language (L2) learners' engagement with AWE-Teacher integrated feedback in a technology-empowered context. *The Asia-Pacific Education Researcher*, 33(4), 1023–1035. doi:10.1007/s40299-024-00877-8
- Cleary, T. J., & Zimmerman, B. J. (2012). A cyclical self-regulatory account of student engagement: Theoretical foundations and applications. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), Handbook of research on student engagement (pp. 237–257). Springer.
- Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology*, 51(1), 171–200. doi:10.1146/annurev.psych.51.1.171
- 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. (2003). Questionnaires in second language research: Construction, administration, and processing. (1st ed.). Routledge.

- Dörnyei, Z., & Taguchi, T. (2009). Questionnaires in second language research: Construction, administration, and processing. (2nd ed.). Routledge.
- Dunn, K., & Iwaniec, J. (2021). Exploring the relationship between second language learning motivation and proficiency. *Studies in Second Language Acquisition*, 44(4), 967–997. doi:10.1017/s0272263121000759
- Ebadi, S., Zandi, M., & Ajabshir, Z. F. (2024). The effect of language learning engagement on L2 motivation over Instagram. *Acta Psychologica*, 250, 104512. doi:10.1016/j.actpsy.2024.104512
- Ene, E., & Yao, J. (2021). How does that make you feel: Students' affective engagement with feedback. Language Teaching Research Quarterly, 25, 66–83. doi:10.32038/ltrq.2021.25.04
- Fan, Y., & Xu, J. (2020). Exploring student engagement with peer feedback on L2 writing. *Journal of Second Language Writing*, 50, 100775. doi:10.1016/j.jslw.2020.100775
- Farsani, M. A., & Aghamohammadi, N. (2021). Exploring students' engagement with peer-and teacher written feedback in an EFL writing course: A multiple case study of Iranian graduate learners. *MEXTESOL Journal*, 45(4), 1–17. doi:10.61871/mj.v45n4-1
- Froiland, J. M., & Davison, M. L. (2016). The longitudinal influences of peers, parents, motivation, and mathematics course-taking on high school math achievement. *Learning and Individual Differences*, 50, 252–259. doi:10.1016/j.lindif.2016.07.012
- Froiland, J. M., & Worrell, F. C. (2016). Intrinsic motivation, learning goals, engagement, and achievement in a diverse high school. *Psychology in the Schools*, 53(3), 321–336. doi:10.1002/pits.21901
- Fu, X., Yang, C., & Zhang, T. (2024). Visibilizing invisible engagement behind students' no-revision operation to written feedback. *Language Teaching Research*. doi:10.1177/13621688241227666
- Gardner, R. C. (1985). Social psychology and second language learning: The role of attitudes and motivation. Edward Arnold.
- Ghelichli, Y., Seyyedrezaei, S., Barani, G., & Mazandarani, O. (2020). The relationship between dimensions of student engagement and language learning motivation among Iranian EFL learners. *International Journal of Foreign Language Teaching and Research*, 8(31), 43–57. https://journals.iau.ir/article_673251. html
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. The Electronic Journal of Business Research Methods, 6(1). doi:10.21427/d7cf7r
- Huang, J. (2023). Engineering ChatGPT prompts for EFL writing classes. *International Journal of TESOL Studies*, 5(4), 73–79. doi:10.58304/ijts.20230405
- Huang, J., & Mizumoto, A. (2024a). Examining the relationship between the L2 motivational self system and technology acceptance model post ChatGPT introduction and utilization. *Computers and Education:* Artificial Intelligence, 100302. doi:10.1016/j.caeai.2024.100302
- Huang, J., & Mizumoto, A. (2024b). The effects of generative AI usage in EFL classrooms on the L2 motivational self system. Education and Information Technologies. doi:10.1007/s10639-024-13071-6
- Huang, J., & Mizumoto, A. (2024c). Examining the effect of generative AI on students' motivation and writing self-efficacy. Digital Applied Linguistics, 1, 102324. doi:10.29140/dal.v1.102324
- Huang, J., Mizumoto, A., & Bailey, B. (2024). Examining the effects of the L2 learning experience on the ideal L2 self and ought-to L2 self in a Japanese university context. *International Journal of Applied Linguistics*. doi:10.1111/ijal.12659
- Huang, J., & Teng, M. F. (2025). Peer feedback and ChatGPT-generated feedback on Japanese EFL students' engagement in a foreign language writing context. *Digital Applied Linguistics*, 2, 102469. doi:10.29140/ dal.v2.102469
- Iwaniec, J., & Khaled, H. Y. (2024). Motivation and engagement in language learning. In A. Cirocki, B. Indrarathne & S. McCulloch (Eds.), Cognitive and educational psychology for TESOL (pp. 131–149). Springer.
- Kanellopoulou, C., & Giannakoulopoulos, A. (2020). Engage and conquer: An online empirical approach into whether intrinsic or extrinsic motivation leads to more enhanced students' engagement. Creative Education, 11(2), 143–165. doi:10.4236/ce.2020.112011
- Katyara, P., Dahri, K. H., & Muhiuddin, G. (2023). Impact of technology on student's engagement in different dimensions: Cognitive, behavioral, reflective and social engagement. Webology, 19(3), 3451–3464.
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). Exploring generative artificial intelligence preparedness among university language instructors: A case study. Computers and Education: Artificial Intelligence, 5, 100156. doi:10.1016/j.caeai.2023.100156

- 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. doi:10.1016/j.caeai.2023.100178
- Lamb, M. (2012). A self system perspective on young adolescents' motivation to learn English in urban and rural settings. *Language Learning*, 62(4), 997–1023. doi:10.1111/j.1467-9922.2012.00719.x
- 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.
- Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 1–14. doi:10. 1080/17501229.2023.2240316
- Liu, Y., Ma, S., & Chen, Y. (2024). The impacts of learning motivation, emotional engagement and psychological capital on academic performance in a blended learning university course. Frontiers in Psychology, 15, 1357936. doi:10.3389/fpsyg.2024.1357936
- Lo, C. K., Hew, K. F., & Jong, M. S. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers and Education*, 219, 105100. doi:10.1016/j.compedu.2024. 105100
- Maričić, M., & Lavicza, Z. (2024). Enhancing student engagement through emerging technology integration in STEAM learning environments. *Education and Information Technologies*. doi:10.1007/s10639-024-12710-2
- Markus, H. R., & Nurius, P. S. (1986). Possible selves. American Psychologist, 41(9), 954–969. doi:10.1037/ 0003-066x.41.9.954
- Martin, A. J., Ginns, P., & Papworth, B. (2017). Motivation and engagement: Same or different? Does it matter? *Learning and Individual Differences*, 55, 150–162. doi:10.1016/j.lindif.2017.03.013
- Mercer, S., & Dörnyei, Z. (2020). Engaging language learners in contemporary classrooms. Cambridge University Press.
- Noels, K. A., Lou, N. M., Lascano, D. I. V., Chaffee, K. E., Dincer, A., Zhang, Y. S. D., & Zhang, X. (2019). Self-determination and motivated engagement in language learning. In M. Lamb, K. Csizér, A. Henry & S. Ryan (Eds.), The Palgrave handbook of motivation for language learning (pp. 95–115). Palgrave Macmillan.
- Papi, M. (2010). The L2 motivational self system, L2 anxiety, and motivated behavior: A structural equation modeling approach. *System*, 38(3), 467–479. doi:10.1016/j.system.2010.06.011
- Park, J., & Ahn, S. (2022). L2 learners' cognitive and behavioral engagement with written corrective feedback. English Teaching, 77(3), 133–152. doi:10.15858/engtea.77.3.202209.133
- Pearson, W. S. (2024). Affective, behavioural, and cognitive engagement with written feedback on second language writing: A systematic methodological review. *Frontiers in Education*, 9, 1285954. doi:10.3389/feduc.2024.1285954
- Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. Christenson, A. Reschly & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 259–282). Springer.
- Purnomo, S. G., & Pahlevi, M. R. (2021). EFL secondary students' affective engagement on teacher written corrective feedback in biography writing class. *Journal of English Language and Education*, 6(2), 27–38. doi:10.31004/jele.v6i2.124
- Rad, H. S., Alipour, R., & Jafarpour, A. (2023). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. *Interactive Learning Environments*, 1–21. doi:10.1080/10494820.2023.2208170
- Reeve, J. (2012). A self-determination theory perspective on student engagement. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), Handbook of research on student engagement (pp. 149–172). Springer.
- Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion*, 28(2), 147–169. doi:10.1023/b:moem.0000032312. 95499.6f
- Renninger, K. A., Ren, Y., & Kern, H. M. (2018). Motivation, engagement, and interest: "In the end, it came down to you and how you think of the problem. In F. Fischer, C. E. Hmelo-Silver, R. G. S. & P. Reimann (Eds.), *International handbook of the learning sciences* (pp. 116–126). Routledge.

- Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3–19). Springer.
- Saeed, S., & Zyngier, D. (2012). How motivation influences student engagement: A qualitative case study. Journal of Education and Learning, 1(2), 252–267. doi:10.5539/jel.v1n2p252
- Schindler, L. A., Burkholder, G. J., Morad, O. A., & Marsh, C. (2017). Computer-based technology and student engagement: A critical review of the literature. *International Journal of Educational Technology in Higher Education*, 14(1), 1–28. doi:10.1186/s41239-017-0063-0
- Schunk, D. H., & Mullen, C. A. (2012). Self-Efficacy as an engaged learner. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), Handbook of research on student engagement (pp. 219–235). Springer.
- Shernoff, D. J., Tonks, S. M., & Anderson, B. (2014). The impact of the learning environment on student engagement in high school classrooms. *Teachers College Record*, 116(13), 166–177. doi:10.1177/016146811411601315
- Sulis, G. (2020). An exploration of fluctuations in motivation and engagement in second language (L2) learning within the lesson and across the academic year Doctoral Thesis]. Lancaster University. doi:10.17635/lancaster/thesis/1060
- Suzuki, M. (2014). Ideal L2 selves of Japanese university freshmen: A change from high school. Soka University Faculty of Engineering Bulletin, 44, 1–8.
- 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.
- Takahashi, C., & Im, S. (2020). Comparing self-determination theory and the L2 motivational self system and their relationships to L2 proficiency. *Studies in Second Language Learning and Teaching*, 10(4), 673–696. doi:10.14746/ssllt.2020.10.4.2
- Teimouri, Y. (2016). L2 selves, emotions, and motivated behaviors. Studies in Second Language Acquisition, 39(4), 681–709. doi:10.1017/s0272263116000243
- Teng, M. F. (2024). "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. doi:10.1016/j.caeai.2024.100270
- Teng, M. F., & Huang, J. (2025). Incorporating ChatGPT for EFL Writing and Its Effects on Writing Engagement. International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT), 15(1), 1–21. doi:10.4018/IJCALLT.367874
- Voelkl, K. E. (2012). School identification. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), Handbook of research on student engagement (pp. 193–218). Springer.
- 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. doi:10.1016/j.chb.2024.108353
- Yang, L., & Zhang, L. J. (2023). Self-regulation and student engagement with feedback: The case of Chinese EFL student writers. *Journal of English for Academic Purposes*, 63, 101226. doi:10.1016/j.jeap.2023.101226
- Yang, T. (2024). Impact of artificial intelligence software on English learning motivation and achievement. SHS Web of Conferences, 193, 02011. doi:10.1051/shsconf/202419302011
- Yashima, T., Nishida, R., & Mizumoto, A. (2017). Influence of learner beliefs and gender on the motivating power of L2 selves. *Modern Language Journal*, 101(4), 691–711. doi:10.1111/modl.12430
- Yu, S., Jiang, L., & Zhou, N. (2020). Investigating what feedback practices contribute to students' writing motivation and engagement in Chinese EFL context: A large scale study. Assessing Writing, 44, 100451. doi:10.1016/j.asw.2020.100451
- 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*. doi:10.1007/s10639-024-12763-3
- Cite this article: Huang, J., & Mizumoto, A. (2025). The role of generative AI in mediating L2MSS and engagement with written feedback in EFL learning: A structural equation modeling approach. *Annual Review of Applied Linguistics*, 45, 131–147. https://doi.org/10.1017/S0267190525000029