


RESEARCH ARTICLE

Interacting with ChatGPT in essay writing: A study of L2 learners' task motivation

Javad Zare 

Kosar University of Bojnord, Iran (javadzare@gmail.com)

Ahmad Al-Issa 

American University of Sharjah, United Arab Emirates (aissa@aus.edu)

Fatemeh Ranjbaran Madiseh 

Sultan Qaboos University, Oman (f.madiseh@squ.edu.om)

Abstract

This study explored the effects of interacting with ChatGPT 4.0 on L2 learners' motivation to write English argumentative essays. Conducted at a public university in a non-English-speaking country, the study had an experimental and mixed-methods design. It utilized both quantitative and qualitative data analyses to inform the development of effective AI-enhanced tailored interventions for teaching L2 essay writing. Overall, the results revealed that interacting with ChatGPT 4.0 had a positive lasting effect on learners' motivation to write argumentative essays in English. However, a decline in their motivation at the delayed post-intervention stage suggested the need to maintain a balance between utilizing ChatGPT as a writing support tool and enhancing their independent writing capabilities. Learners attributed the increase in their motivation to several factors, including their perceived improvement in essay writing skills, the supportive learning environment created by ChatGPT as a tutor, positive interactions with it, and the development of meta-cognitive awareness by addressing their specific writing issues. The study highlights the potential of AI-based tools in enhancing L2 learners' motivation in English classrooms.

Keywords: generative AI; ChatGPT; task motivation; essay writing; L2 learners

1. Introduction

Motivation is an extensively studied concept in second language acquisition (SLA) (e.g. Lamb, 2016; Ushioda, 2008). Dörnyei and Ushioda (2013) define it as “what moves a person to make certain choices, to engage in action, to expend effort and persist in action” (p. 3). Second language (L2) motivation is complex and dynamic and can significantly impact the extent to which a learner will engage with and succeed in learning (Dörnyei, 2002; Dörnyei, MacIntyre & Henry, 2015). It can vary depending on the setting and the type of language activity, hence the concept of task motivation (Jülkunen, 1989). Task motivation is the motive that arises in response to specific task features (Poupore, 2013). Dörnyei (2019) highlights that task motivation is a multifaceted, dynamic, and fluid concept that involves the interplay between various factors, including learner characteristics, learning context, and the task. Therefore, it is essential to study motivation in relation to different task features.

Cite this article: Zare, J., Al-Issa, A. & Ranjbaran Madiseh, F. (2025). Interacting with ChatGPT in essay writing: A study of L2 learners' task motivation. *ReCALL* FirstView, 1–18. <https://doi.org/10.1017/S0958344025000035>

© The Author(s), 2025. Published by Cambridge University Press on behalf of EUROCALL, the European Association for Computer-Assisted Language Learning. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Writing English essays is demanding for L2 learners, due to the various cognitive, linguistic, and motivational resources it requires (Kormos, 2023). This highlights the need to provide personalized instruction and feedback to address each learner's unique needs and challenges in L2 essay writing (Ferris & Hedgcock, 2023). Therefore, it is essential to implement interventions that are aimed at fostering L2 learners' motivation to write English essays.

At the same time, advances in artificial intelligence (AI) have led to the development of chatbots that have been found to be highly beneficial in L2 development (Zhang & Huang, 2024). Chatbots can provide personalized instruction in various aspects of L2 development, including essay writing (e.g. Guo, Wang & Chu, 2022; Su, Lin & Lai, 2023), although the feedback they generate in many cases may be generic and not tailored to address specific learner profiles. ChatGPT is one such tool. As a state-of-the-art generative AI chatbot, ChatGPT (introduced in 2022 by OpenAI) utilizes deep learning techniques to produce human-like text. It is designed to engage in conversations with users and provide contextually relevant responses to prompts. In L2 education, ChatGPT can be a valuable tool for developing essay writing skills (Guo, Wang & Chu, 2022; Su, Lin & Lai, 2023). Its interactive feedback and adaptive responses may make it an excellent fit for personalized learning experiences, aligning well with the principles of learner-centered instruction in SLA (Barrot, 2023), which stresses the uniqueness of learners and the individual differences they bring to the learning process (Cohen & Dörnyei, 2002). Yet students need to be able to critically assess and make use of the feedback they receive from AI chatbots. This requires the ability to determine the quality, relevance, and applicability of the feedback they receive.

Despite its potential benefits, research on its applications for L2 development, especially essay writing, is very limited (e.g. Guo, Wang & Chu, 2022; Su, Lin & Lai, 2023). Aiming to fill this gap, in this study we explored the effects of interacting with ChatGPT 4.0 on L2 learners' motivation to write English essays. With its experimental and mixed-methods design, the study sought to inform the development of effective AI-enhanced tailored interventions for teaching L2 essay writing.

2. Review of the literature

2.1 Task motivation in SLA: An overview

Motivation is widely recognized as a crucial factor for success in L2 learning. It is generally seen as the driving force that propels the acquisition process. Task motivation, introduced by Jülkunen in 1989, refers to how motivated learners are when completing a language activity. It encompasses state-like appraisals of tasks that involve task-specific cognitions and feelings evoked while performing the task (Seegers & Boekaerts, 1993). Task motivation is a complex concept that involves a dynamic interplay between various learner-related, context-specific, and task-driven factors (Dörnyei, 2019). Learner-specific factors include personality traits, competence, and learning styles. Contextual factors refer to the impact of the teacher, group dynamics, and other situational variables. Finally, task features include the structure, content, and outcome of the task itself. Therefore, it is crucial to examine task motivation and its connection with various factors.

There have been various studies investigating task motivation within different L2 settings. Research has explored how the characteristics of language learning tasks can impact students' motivation to complete them (e.g. Dörnyei & Murphey, 2003; Kormos & Préfontaine, 2017; Mozgalina, 2015; Poupore, 2013; Zare & Aghajani Delavar, 2024). Studies show that selecting content that is personally relevant to learners and allows for self-expression can enhance their task motivation (Kormos & Wilby, 2019). Additionally, giving learners the freedom to create their own task content can also increase their task motivation, but it is crucial to strike a balance between free choice and teacher direction (Mozgalina, 2015). Furthermore, tasks must also be tailored to the appropriate level of difficulty to keep motivation high (Kormos & Préfontaine, 2017; Poupore, 2013). Peers and group dynamics also play a significant role in task motivation, especially for tasks

that require interaction with peers. Motivation for such tasks is collectively constructed and relies heavily on the functioning of both small groups and the entire class (Kormos & Wilby, 2019). Finally, students are more motivated when they are actively involved in their learning through discovery-oriented activities (Zare & Aghajani Delavar, 2024).

Task motivation can be understood through the lens of self-determination theory (SDT), self-regulated learning (SRL), and learner-centered instruction. SDT suggests that motivation can be enhanced when individuals perceive a sense of control over their actions (autonomy), belief in their abilities (competence), and connection to others (relatedness) (Ryan & Deci, 2017). Applied to the L2 learning context, SDT illuminates how task motivation can be influenced by the degree to which the learning environment caters to these three needs. That is, when a task allows learners to have a sense of control, challenges them at the appropriate level, and promotes a sense of connection to others, it can boost their intrinsic motivation. Conversely, tasks that undermine these needs may foster extrinsic motivation, where the drive to complete a task is based on external rewards or pressures. This framework is particularly useful for designing tasks in L2 classrooms. For example, tasks that offer choice, promote student agency, and align with learners' interests and goals are more likely to satisfy autonomy, thereby enhancing intrinsic motivation (Zare & Aghajani Delavar, 2024). Additionally, tasks that scaffold learners toward achievable challenges encourage competence without causing anxiety (Kormos & Préfontaine, 2017). Finally, tasks that are socially relevant or involve cooperation can foster a sense of relatedness and increase learner's motivation (Kormos & Wilby, 2019).

SRL, on the other hand, concerns the extent to which learners are metacognitively, motivationally, and behaviorally active in their learning process (Zimmerman, 2011). SRL describes how learners plan, monitor, and evaluate their learning goals and outcomes (Zimmerman, 2008). In L2 learning contexts, learners who exhibit high levels of SRL are more likely to set specific objectives, use effective strategies, monitor their progress, and adjust their approaches (Fukuda, 2018). SRL shares commonalities with SDT in autonomy. L2 tasks that promote SRL not only enable learners to take ownership of their learning but also empower them to confront challenges by using adaptive strategies (Weinstein, Acee & Jung, 2011). This, in turn, enhances their motivation, as they experience a sense of personal growth and success while working toward improving their language proficiency (Chen, Chen & Yang, 2019).

Finally, learner-centered instruction places the learner at the heart of instruction to promote a more personalized and meaningful learning experience (Richards & Schmidt, 2010). In this framework, teachers must acknowledge and consider the diversity of learners in learning to create an effective learning environment (Skehan, 1989). One such difference that learners bring to the language classroom is their motivation (Dörnyei, 2005). Additionally, this framework encourages learners to take an active role in their learning process. Hence, tasks in learner-centered instruction aim to teach learners how to learn, not simply to practice language skills. Engaging in reflective activities such as setting goals, monitoring progress, applying strategies, and self-assessment is crucial for becoming a self-sufficient, strategic, and motivated learner.

2.2 Argumentative essay writing, ChatGPT, and task motivation

The ability to write effectively in academic contexts has become an essential skill, closely linked to academic success and professional growth (Wang, Derakhshan, Pan & Ghiasvand, 2023). Academic writing is by no means an easy task, especially for those who are navigating the complexities of a second language. It requires a thorough understanding of subtle linguistic features, as well as an awareness of the specific structural and rhetorical patterns that academic discourses involve (Fang, 2021). Among various forms of academic discourse, argumentative essays are considered a cornerstone and predictor of academic literacy and success (Hartwell & Aull, 2022). Crafting an argumentative essay goes beyond mere linguistic ability; it involves advanced cognitive skills such as reasoning, gathering, comparing, analyzing, synthesizing, and connecting ideas to form a clear and

well-founded position, as well as developing sound arguments supported by evidence (McNeill & Krajcik, 2009). These demands can be overwhelming and cause significant stress to L2 learners, which may hinder their learning (Wei, Zhang & Zhang, 2020). On the other hand, language educators face the difficult task of guiding students through the complex process of developing an argumentative essay. The large number of learners in most classes, as well as their diverse writing needs, adds to the complexity of teaching how to write English argumentative essays for several reasons. First, larger class sizes limit individualized attention and feedback. Second, students have varying proficiency levels, learning styles, and prior knowledge in writing argumentative essays, making personalized instruction challenging in a classroom with limited resources and time constraints. This can put a strain on resources and make it difficult to provide the individualized attention each student needs (Guo, Wang & Chu, 2022).

Yet it is extremely important to provide personalized feedback to each student. Such tailored guidance is essential for helping students improve their linguistic and cognitive skills, and ensuring that they receive the necessary support and enrichment to succeed in mastering English essay writing (Olsen & Hunnes, 2024). Personalized feedback not only helps students identify their strengths and weaknesses in writing but also gives them a sense of direction and purpose in their learning process (Weaver, 2006). This type of feedback can be very helpful, as it provides a deeper understanding that generic comments cannot offer (Hattie & Timperley, 2007). It can address the subtle linguistic nuances, conceptual intricacies, and logical subtleties that students may face when writing English argumentative essays.

In response to this issue, educators have recognized the importance of using technological advancements, such as AI-based chatbots, to provide personalized learning experiences to learners (e.g. Barrot, 2023; Guo, Wang & Chu, 2022; Su *et al.*, 2023). Such tools can provide immediate, situation-specific, and adaptive feedback (Rudolph, Tan & Tan, 2023). One such tool is ChatGPT, accessible via desktops and mobile devices. ChatGPT is an AI chatbot, powered by advanced computing techniques and capable of generating situation-specific human-like responses. ChatGPT can help improve the quality of students' writing by proofreading, editing, and suggesting corrections for their errors in grammar, syntax, and textual devices (Geher, 2023; McMurtrie, 2022). In argumentative essay writing tasks, it can facilitate the various steps of writing such essays, including developing ideas, editing, proofreading, and reflecting (Su *et al.*, 2023). Aside from undermining learners' creativity, critical thinking, and integrity, ChatGPT has a lot of potential as a writing tutor or assistant (Barrot, 2023). Therefore, it is important to investigate how it can enhance different aspects of L2 learning, including essay writing. To this end, this study investigated if, to what extent, and how interacting with ChatGPT affected learners' motivation to write English argumentative essays. More specifically, the study aimed to answer the following research questions:

1. Whether and to what extent does interacting with ChatGPT 4.0 influence L2 learners' motivation to write English essays?
2. What do participants report about their experience with ChatGPT 4.0?

3. Methods

3.1 Context and design

This experimental research was conducted at the English department of a public university in Iran, where English is taught as a foreign language. In such settings, English is mainly used in educational contexts and when people encounter foreign visitors. The study's context is an ideal place to examine the dynamics of EFL education.

The study was conducted using a mixed-methods design, namely sequential explanatory, including both quantitative (obtained from motivation questionnaires) and qualitative (participant interviews) analyses. Quantitative analysis helped answer the first research question by providing statistical evidence of the extent to which interacting with ChatGPT influenced

motivation, while qualitative analysis gave insights into the reasons behind these changes, hence addressing the second research question.

A pre-test/post-test control group experimental design was adopted as well. Therefore, participants were randomly assigned to either the comparison or the intervention group, with the former engaging in traditional pair work writing tasks and the latter interacting with ChatGPT 4.0. To evaluate the effects of the intervention on learners' motivation for writing English essays, we took measures at three distinct time points: before the intervention (pre-testing), immediately following the intervention (immediate post-testing), and one month after the intervention (delayed post-testing). The rationale for adopting a pre-test/post-test control group experimental design in the study was that it allowed for comparing the effect of interacting with ChatGPT on L2 learners' motivation to write English essays. By including a control group that did not interact with ChatGPT, any post-intervention changes in the experimental group learners' motivation could be attributed to the intervention rather than to external factors. Additionally, the pre-test/post-test design allowed for measuring changes in motivation over time, providing a clearer picture of the impact of interacting with ChatGPT on L2 learners' motivation to write English essays. Including both pre-test and post-tests as well as control and experimental groups were aimed to control for the effects of confounding variables.

Altogether, the combination of a pre-test/post-test control group experimental design and a sequential explanatory mixed-methods design helped us provide a deeper understanding of the mechanisms through which interacting with ChatGPT influenced L2 learners' motivation to write English essays.

3.2 Participants

The study's participants consisted of 69 university students, selected from around 200 students studying English at the English department of a public university in Iran. They were all native speakers of Persian. The recruitment process involved selecting participants in a way that ensured representation of different genders, age groups, and academic years. To ensure a consistent level of motivation among participants and to mitigate the potential impact of external factors, we established additional criteria for recruitment. These included willingness to participate in the study, a comparable level of proficiency (both in general English and writing), as well as task motivation. Hence, they had comparable levels of English proficiency (B1), writing skills, and motivation.

Participant selection was done through a stratified random sampling technique. This involved dividing the English students into subgroups, based on criteria such as age, academic year, and gender. The purpose was to make sure that each subgroup was represented in the final sample. After the students were divided into these subgroups, participants were randomly chosen from each one to form the final sample. This approach was used to prevent potential biases and to ensure that the group of participants accurately represented the diverse characteristics of the larger student population.

Regarding the measurement and control of participants' baseline competencies, we assessed the students' English proficiency level (B1), writing skills, and task motivation during the recruitment process. These criteria were used to ensure that all participants had comparable competencies at the start of the study, thus providing a more controlled and standardized basis for the research outcomes. It is important to note that given the design of the study, we needed to recruit homogeneous groups. Hence, we established a baseline for general proficiency, writing, and motivation at the beginning of the study. In experimental research designs, recruiting homogeneous groups is crucial for establishing a baseline, as it minimizes variability in key characteristics among participants. By ensuring that participants shared similar attributes, such as proficiency level, writing, and motivation, we could isolate the effect of the independent variable on the dependent variable more effectively. This controlled environment may enhance the internal

validity of the study and lead to more reliable conclusions regarding the causal relationships observed. Moreover, homogeneous groups help reduce the risk of confounding factors influencing the outcomes, allowing for clearer interpretations of the intervention's effects.

The final sample included 14 males (20.29%) and 55 females (79.71%); six were second-year students (8.70%); 22 were third-year students (31.88%), and 41 were fourth-year students (59.42%). Their age varied from 19 to 28 ($M = 21.50$, $SD = 2.24$).

Participants gave their informed consent, acknowledging their understanding of the study's purpose and their rights as volunteers, including confidentiality and the ability to withdraw without consequences.

3.3 Measures

3.3.1 Oxford Online Placement Test

The Oxford Online Placement Test was used to assess the participants' general English proficiency and ensure that it did not influence the learners' motivation to write. This test assesses grammar, vocabulary, and listening comprehension, reporting results based on the CEFR (Pre-A1 to C2).

3.3.2 TOEFL iBT independent writing tasks

To ensure that learners were comparable in their writing skills before the intervention, we evaluated their essay writing skills with a TOEFL iBT independent writing task. Widely established as a reliable and valid measure of one's writing skills, a TOEFL iBT independent writing task typically asks learners to develop an English essay in response to a prompt in 30 minutes. Three tasks were used in this investigation: one before the intervention and two after it. The two subsequent tasks were used as a reference for completing the task motivation questionnaire.

3.3.3 Task motivation questionnaire

To evaluate learners' task motivation, we used the task motivation questionnaire, developed by Ma (2009). The questionnaire included 24 items on a 6-point Likert scale. The subscales included perceived choice, perceived competence, relatedness, intrinsic motivation, identified regulation, external regulation, amotivation, and intention to proceed, each consisting of three items. The questionnaire included three sections: the first gathered demographics, the second presented items, and the third collected contact details in case learners wanted to participate in the follow-up interview.

The questionnaire was validated in a study by Zare and Aghajani Delavar (2024). The results of principal components analysis with varimax rotation indicated that the sample was sufficient for factor analysis, with a Kaiser–Meyer–Olkin measure of sampling adequacy at 0.813 and a significant Bartlett's test of sphericity. The questionnaire also demonstrated good reliability with a Cronbach's alpha coefficient of 0.74, indicating acceptable internal consistency. Furthermore, the questionnaire's accuracy, clarity, and thoroughness were verified by an expert.

The questionnaire was administered (through Google Forms) three times during the study. The first time was before the intervention, to ensure that learners had similar task motivation levels. The second and third times were immediately after the intervention and with a one-month interval, respectively, to identify any differences between learners in the comparison and intervention groups.

3.3.4 Interviews

Online interviews were held with voluntary learners from both groups (eight from each group) to gain a better understanding of how interacting with ChatGPT influenced learners' motivation in writing English essays. These interviews were in Persian, the learners' native language, and were

semi-structured, allowing learners to share their experiences with ChatGPT while writing English essays. The interview questions covered topics such as learners' increased/decreased motivation, specific aspects of ChatGPT that influenced their motivation to write English essays, and any other relevant issues they experienced during the intervention. The interviews were conducted after the intervention and lasted approximately 20 minutes on average.

3.3.5 ChatGPT

ChatGPT 4.0 is a version of the Generative Pre-trained Transformer series developed by OpenAI. It serves as an AI conversational agent that can comprehend and generate natural language responses. Using advanced language processing capabilities, it delivers coherent and relevant output tailored to user prompts. Its versatility positions ChatGPT 4.0 as an invaluable tool when writing English essays. Participants accessed it through the website <https://start.chatgot.io/>.

3.4 Procedures

Pre-testing involved taking the Oxford Online Placement Test, completing a TOEFL iBT independent writing task, and filling out the task motivation questionnaire. Pre-testing aimed to mitigate the impact of differences in students' English proficiency, writing skills, and motivation on their post-intervention performance. Based on their scores, we selected 69 students with similar English proficiency, writing skills, and motivation using stratified sampling to ensure a balanced distribution of demographic and academic variables such as gender, age, and academic year between the comparison and intervention groups.

The intervention and control intervention centered on teaching English argumentative essay writing. Both interventions consisted of 12 twice-weekly sessions, taught by one English teacher, following the "three Ps" methodology (presentation, practice, and production). During the presentation phase, the instructor laid out the framework for developing an English argumentative essay. This phase began by presenting the framework and key components of developing an English argumentative essay. It included explanations, examples, and models to help students understand the structure and requirements of an English argumentative essay. Following the presentation phase, students engaged in practice activities to reinforce their understanding of the concepts covered. These practice exercises included analyzing sample essays, identifying key elements of argumentation, and practicing with structured writing exercises. Through these activities, students got hands-on practice applying the concepts presented in the earlier phase. The final phase was production, where students were engaged in actually writing argumentative essays. In this phase, students applied the knowledge and skills they had gained from the presentation and practice phases to create their own original essays. The production phase was where students demonstrated their ability to develop arguments, structure their ideas cohesively, and effectively communicate their thoughts in writing.

The difference between the comparison and intervention groups was that, during the production phase, the comparison group collaborated with peers, whereas the intervention group interacted with ChatGPT 4.0 on their smartphones (through accessing <https://start.chatgot.io/>) to refine ideas, edit, and proofread their essays. Learners were given clear instructions regarding the use of ChatGPT – what they were allowed to seek help for and what they were expected to accomplish independently. The teacher closely monitored intervention group students to ensure they only used ChatGPT for (1) refining their ideas; (2) commenting on their initial drafts; and (3) suggesting corrections on vocabulary, grammar, and the use of textual devices. The teacher prohibited using ChatGPT for producing the final essay. This involved conducting in-session observations and debriefs. Such measures ensured that learners used ChatGPT only as a supplementary and support resource. Furthermore, as the students were unfamiliar with ChatGPT, three prompts were prepared in advance for them to use when interacting with

ChatGPT in writing English argumentative essays. These prompts focused on refining ideas; commenting on initial drafts; and suggesting corrections on vocabulary, grammar, and textual devices. For the comparison group, however, the production phase activities were done in pairs. Instead of using ChatGPT, they received feedback and support from their peers for refining ideas; commenting on initial drafts; and suggesting corrections on vocabulary, grammar, and textual devices when writing English argumentative essays.

Following the intervention, both groups took the task motivation questionnaire twice: immediately and a month later, as immediate and delayed post-tests. They were also provided with a TOEFL writing task for reference while answering the task motivation questionnaire items.

The rubrics used to assess the TOEFL writing tasks were those of TOEFL iBT independent writing. These include effectively addressing the topic and task, ensuring the essay is well organized and developed with clear explanations and examples, exhibiting unity, progression, and coherence in the essay structure, and demonstrating consistent proficiency in language usage with varied syntax, appropriate vocabulary, and idiomatic expressions, despite the possibility of minor lexical or grammatical errors.

3.5 Data analysis

The quantitative data analysis involved parametric tests, as the pre-test, immediate, and delayed post-test results verified normality ($p = 0.108, 0.245, 0.323 > 0.05$). Subsequently, we ran three separate independent-samples *t*-tests, one for each evaluation stage, using SPSS Version 26. These tests allowed for the comparison of mean differences between groups at each evaluation stage. Before this, the assumptions dealing with such statistical tests, including independence of observations, normality, and homogeneity of the variances, were tested and confirmed. Additionally, to assess significant changes over the study's timeline – from pre-intervention to immediate and delayed post-intervention – we conducted repeated measures ANOVAs for both the comparison and intervention groups. The assumptions for repeated measures ANOVA, including but not limited to sphericity, were also tested and verified.

To analyze the qualitative data, we transcribed, translated (into English), and reviewed participants' responses for possible errors or inconsistencies. To reduce bias, we employed an emergent coding approach, which involved deriving codes and themes from a thorough and iterative process. This approach was based on Braun and Clarke's (2006) model, which included five steps: reading the data repeatedly to fully grasp their scope; creating initial codes; grouping codes into broader themes; refining, defining, and naming the themes; and finally compiling a report containing anonymous excerpts.

To make sure the analyses were trustworthy, we took various measures, including credibility, transferability, confirmability, and dependability (Lincoln & Guba, 1985). Credibility was supported through data triangulation, employing both quantitative and qualitative measures, while transferability was enhanced by compiling a report detailing the study's setting, design, participants, measures, and derived themes. Confirmability was addressed through independent coding by two coders, with their inter-coder reliability computed using Cronbach's alpha ($\alpha = 0.88$). Additionally, a coder (foreign coder) from a different cultural context was brought in to code and validate a portion of the data ($\alpha = 0.85$), thus reinforcing the confirmability of the analysis. The foreign coder was fluent in English but came from a different cultural background and linguistic community than the primary researchers. He also had a strong background in SLA research and qualitative research methods and was selected because of his expertise and experience in coding and analyzing qualitative data. He was involved in the coding process to provide a different perspective on the interpretation of data, thereby enhancing the confirmability of our qualitative analysis. Recruiting the foreign coder also enabled us to address researcher positionality by maintaining both emic (insider) and etic (outsider) perspectives. Dependability

Table 1. Task Motivation Questionnaire results

	Group	<i>N</i>	<i>M</i>	<i>SD</i>	Std. error mean
Pre-test	Comparison	38	90.73	19.22	3.11
	Intervention	31	91.67	21.78	3.91
Immediate	Comparison	38	94.89	17.18	2.78
	Intervention	31	107.38	17.42	3.12
Delayed	Comparison	38	96.89	18.32	2.97
	Intervention	31	103.51	14.08	2.53

Note. *N* = number; *M* = mean; *SD* = standard deviation; Std. error mean = standard error of mean.

was improved by conducting member checking, whereby five participants reviewed the derived themes to confirm their alignment with their opinions (Creswell, 2014).

3.6 Ethical considerations

Throughout our analysis, we prioritized willingness to participate in the study, informed consent, confidentiality, and the right to withdraw without consequences. The excerpts use pseudonyms.

4. Results

4.1 The results of quantitative analyses

The first research question aimed to explore the impact and extent of interacting with ChatGPT 4.0 on L2 learners' motivation to write English essays. Table 1 presents descriptive statistics, providing details regarding the learners' motivation to write English essays at the three time points of the study.

As Table 1 displays, at the pre-test stage, the comparison ($M = 90.73$, $SD = 19.22$) and intervention groups ($M = 91.67$, $SD = 21.78$) showed comparable motivation levels. Following the intervention, the intervention group displayed a substantial increase in motivation at the immediate post-test stage ($M = 107.38$, $SD = 17.42$). The comparison group also showed improvement ($M = 94.89$, $SD = 17.18$), although not as substantial as that of the intervention group. Moving to the delayed post-test, the intervention group's mean score ($M = 103.51$, $SD = 14.08$) slightly decreased from the immediate post-test but remained high, indicating continued improvement over time. This could be attributed to the novelty effect of ChatGPT or external factors influencing the participants during the delayed post-test period. Conversely, the comparison group ($M = 96.89$, $SD = 18.32$) displayed a small rise in motivation.

To determine the statistical significance of the observed differences, we performed separate independent-samples *t*-tests. As Table 2 shows, there was no significant difference between the comparison and intervention groups in their motivation to write English essays before the intervention ($p = 0.850$, $0.852 > 0.05$). The effect size also indicated a negligible effect, confirming that the groups had similar starting points in terms of motivation to write English essays (Cohen's $d = 0.046$), as per Plonsky and Oswald (2014).

Immediately after the intervention, the intervention group outperformed the comparison group significantly ($p = 0.004 < 0.05$), with an effect size of 0.722. This indicates a moderate impact of the intervention, with the intervention group experiencing a notably greater increase in motivation to write English essays than the comparison group (Plonsky & Oswald, 2014).

However, at the delayed post-test stage, the difference was not statistically significant ($p = 0.103$, $0.095 > 0.05$), with a medium effect size of 0.4 (Plonsky & Oswald, 2014). This

Table 2. Independent-samples *t*-tests

		Levene's test for equality of variances		t-test for equality of means					95% CI of the difference	
		<i>F</i>	Sig.	<i>t</i>	<i>df</i>	Sig. (2-tailed)	Mean difference	Std. error difference	Lower	Upper
Pre-test	Equal variances assumed	.062	.803	-.190	67	.850	-.94	4.94	-10.80	8.92
	Equal variances not assumed			-.188	60.44	.852	-.94	5.00	-10.94	9.06
Immediate	Equal variances assumed	.104	.749	-2.98	67	.004	-12.49	4.18	-20.84	-4.13
	Equal variances not assumed			-2.98	63.89	.004	-12.49	4.19	-20.86	-4.11
Delayed	Equal variances assumed	.938	.336	-1.65	67	.103	-6.62	4.00	-14.62	1.38
	Equal variances not assumed			-1.69	66.79	.095	-6.62	3.90	-14.41	1.17

Note. CI = confidence interval.

Table 3. Tests of within-subjects effects for the comparison group

		Measure: Motivation					
Source		Type III sum of squares	<i>df</i>	Mean square	<i>F</i>	Sig.	Partial eta squared
Time	Sphericity assumed	749.96	2	374.98	1.271	.287	.03
	Greenhouse–Geisser	749.96	1.86	401.83	1.271	.286	.03
	Huynh–Feldt	749.96	1.96	382.30	1.271	.286	.03
	Lower bound	749.96	1.00	749.96	1.271	.267	.03
Error (Time)	Sphericity assumed	21828.70	74	294.98			
	Greenhouse–Geisser	21828.70	69.05	316.10			
	Huynh–Feldt	21828.70	72.58	300.74			
	Lower bound	21828.70	37.00	589.96			

Table 4. Pairwise comparisons for the comparison group

		Measure: Motivation					95% CI for difference	
(I) Time	(J) Time	Mean difference (I–J)	Std. error	Sig.	Lower bound	Upper bound		
1	2	–4.15	4.00	.918	–14.20	5.88		
	3	–6.15	4.34	.493	–17.04	4.72		
2	1	4.15	4.00	.918	–5.88	14.20		
	3	–2.00	3.42	1.000	–10.57	6.57		
3	1	6.15	4.34	.493	–4.72	17.04		
	2	2.00	3.42	1.000	–6.57	10.57		

Note. CI = confidence interval.

suggests that the increase in the intervention group's motivation did not maintain statistical significance over time.

We performed two separate repeated measures ANOVAs to see if the differences in each group's motivation across pre-intervention, immediate, and delayed post-intervention were significant. The results are reported in Tables 3–6. As Table 3 shows, no difference was found for the comparison group learners, $F(1.86, 69.05) = 1.271$, $p = 0.286 > 0.05$.

According to the post hoc analysis, the comparison group's motivation did not vary significantly across pre-intervention, immediate, and delayed post-intervention stages ($p = 0.918, 0.493, 1.000 > 0.05$).

Table 5 indicates a statistically significant difference for the intervention group, $F(1.47, 44.10) = 6.689$, $p = 0.006 < 0.05$.

Table 5. Tests of within-subjects effects for the intervention group

		Measure: Motivation					
Source		Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared
Time	Sphericity assumed	4153.31	2	2076.65	6.689	.002	.18
	Greenhouse–Geisser	4153.31	1.47	2825.41	6.689	.006	.18
	Huynh–Feldt	4153.31	1.52	2719.65	6.689	.006	.18
	Lower bound	4153.31	1.00	4153.31	6.689	.015	.18
Error (Time)	Sphericity assumed	18627.35	60	310.45			
	Greenhouse–Geisser	18627.35	44.10	422.39			
	Huynh–Feldt	18627.35	45.81	406.58			
	Lower bound	18627.35	30.00	620.91			

Table 6. Pairwise comparisons for the intervention group

		Measure: Motivation					
(I) Time	(J) Time	Mean difference (I–J)	Std. error	Sig.	95% CI for difference		
					Lower bound	Upper bound	
1	2	–15.71	5.65	.028	–30.04	–1.37	
	3	–11.83	3.91	.015	–21.76	–1.90	
2	1	15.71	5.65	.028	1.37	30.04	
	3	3.87	3.57	.863	–5.19	12.93	
3	1	11.83	3.91	.015	1.90	21.76	
	2	–3.87	3.57	.863	–12.93	5.19	

Note. CI = confidence interval.

As shown in Table 6, for the intervention group, the post hoc analysis indicated a significant variation in learners' motivation between the pre-intervention stage and the immediate and delayed post-intervention stages ($p = 0.028$, $0.015 < 0.05$). However, they did not exhibit any significant variations between the immediate and delayed post-intervention stages ($p = 1.000 > 0.05$). These results suggest that interacting with ChatGPT 4.0 when writing English essays positively impacted learners' motivation to write such essays.

4.2 The results of qualitative analyses

To investigate how interacting with ChatGPT 4.0 influenced learners' motivation to write English essays, we analyzed their responses to the interviews and identified four key themes: (1) progress, (2) support, (3) enjoyment, and (4) individualized feedback. Table 7 displays the frequency and percentage of these themes.

As can be seen, *progress* was the most frequently mentioned theme (38.89%). Concerning this theme, Jack and Mary expressed the following:

Table 7. Thematic analysis of the interviews

Themes	<i>N</i>	%
Progress	7	38.89
Support	4	22.22
Enjoyment	4	22.22
Individualized feedback	3	16.67
Total	18	100

- (1) When I first started using ChatGPT, I had difficulty organizing my thoughts. But, after a couple of weeks, I can structure my essays much better. I think the suggestions and practice taught me how to do it. (Jack, a 24-year-old third-year English learner)
- (2) I guess I know more English words now. Each time ChatGPT gave me feedback, I learned new words and how to use them in context. I feel like my writing has become much better now. (Mary, a 24-year-old third-year English learner)

Jack's observation indicates that interacting with ChatGPT led to improvement in writing English essays in general and structuring essays in particular. Similarly, Mary's experience suggests that the feedback she received from ChatGPT has expanded her English vocabulary, which suggests an enhancement in her linguistic competence. Both excerpts point to a perceived sense of enhancement in essay writing skills or competence in English.

The next theme that emerged from the analysis was *support*. The following statements are from Susan and David about this theme:

- (3) It felt like having a tutor available all the time, someone I looked up to when I needed. Whenever I had a problem with writing, I would ask for suggestions, and it was right there, offering something helpful. (Susan, a 24-year-old fourth-year English learner)
- (4) For me, it was the instant feedback. Getting immediate help with grammar and everything has made a huge difference for me. I didn't have to wait for a teacher to check my draft; the support was always there when I needed it. (David, a 20-year-old third-year English learner)

The statements indicate that Susan and David both perceived ChatGPT as a reliable source of guidance for their writing. Susan viewed ChatGPT as a tutor, while David emphasized the value of the immediate feedback he received from it in grammar and writing. Both statements highlight the role of consistent feedback from a reliable source in creating a supportive learning environment, which in turn leads to greater satisfaction and persistence in learning.

Another recurring theme in the learners' responses was *enjoyment*. Nancy and Sarah shared their views regarding this theme in the following two excerpts:

- (5) Writing assignments has always been a nightmare for me. But now, I think it is kind of fun. I enjoy experimenting with different ways to say things. You write something and you get a response. It's like a game. That's why I enjoy experimenting with it. With ChatGPT, writing has become fun. (Nancy, a 23-year-old second-year English learner)

- (6) Seeing my writing improve has been really exciting. The more I interact with it, the more I want to write just to see how much better I can get. It was a pleasant experience. I wish we had more free tokens every day. (Sarah, a 26-year-old fourth-year English learner)

Nancy's experience with ChatGPT has transformed her perception of writing from a "nightmare" to an enjoyable activity. She attributed this shift to the interactive and responsive nature of ChatGPT, likening it to a game. Sarah's eagerness to write more and see further improvement in her skills shows increased intrinsic motivation, due to the positive interactions she had with this AI tool. Her statement also highlights the importance of tracking one's progress in learning as an intrinsic motivator.

Individualized feedback was the last frequently mentioned theme by the interviewees. In this regard, Paul and Leila highlighted the following:

- (7) What I liked most about it was the feedback which was related to my problems in English. I remember it made me realize how I used too much passive voice or unnecessary conjunctions, which I hadn't noticed in my writing. So, I think part of the improvement was because it told me what I needed to work on. (Paul, a 25-year-old third-year English learner)
- (8) The comments it gave me were very helpful. I used to write long sentences, which made it hard to follow the argument. Points like this helped me focus on my problems and write better. (Leila, a 20-year-old second-year English learner)

Paul highlights the crucial role ChatGPT played in improving his essay writing skills by making him aware of the specific issues in his writing style, such as excessive use of passive voice and conjunctions. Similarly, Leila's experience with ChatGPT suggests that the tool helped her correct certain problems in her writing, such as long sentence construction. Both learners' insights suggest that using such AI-powered tools can enhance their meta-cognitive awareness, which in turn improves their essay writing skills. This reinforces the personalized aspect of ChatGPT's feedback and responses.

5. Discussion

The study explored whether, to what extent, and how interacting with ChatGPT 4.0 influenced L2 learners' motivation to write English essays. Altogether, the results suggested that interacting with ChatGPT 4.0 when writing English essays positively affected learners' motivation to write such essays. Yet the results pointed to a decrease in the intervention group learners' motivation at the delayed post-intervention stage of the analysis, which may be attributed to their temporary dependence on ChatGPT as a technological aid, which later led to reduced independent effort. Learners attributed their increased motivation to their perceived sense of improvement in essay writing skills, the supportive learning environment ChatGPT created by serving as a tutor and providing constant feedback, the positive interactions they had with it, and the meta-cognitive awareness it developed in them by reminding them of their specific problems in writing English essays.

The findings of the study align with several prominent theories in education, including the SDT, SRL, and learner-centered instruction. SDT emphasizes the importance of self-regulation in learning behaviors and categorizes them based on their level of self-regulation or self-determination. According to SDT, behaviors can range from the least determined to the most self-determined, depending on their level of self-direction. Behaviors that are highly self-determined foster intrinsic motivation, where an individual engages in an activity purely for the enjoyment or satisfaction it gives. Certain fundamental psychological needs such as autonomy,

competence, and relatedness are necessary for intrinsic motivation (Ryan & Deci, 2017). Fulfillment of these needs is highlighted in the interviewees' responses. For example, Jack's and Mary's expressions of "progress" reflect an increased sense of competence. Susan's and David's experiences of "support" fulfill the need for relatedness with ChatGPT as a tutor figure. Autonomy is evident in Jack's and Mary's active involvement in the learning process by taking ownership of their learning. Additionally, Sarah's enthusiasm to write more and observe improvement and Paul's realization and subsequent correction of his overuse of passive voice and conjunctions indicate that they have taken control of their learning, setting personal objectives and monitoring their progress.

The findings can also be interpreted through the lens of SRL. Learning, according to SRL, is a self-directed process, where learners regulate their thoughts, emotions, and actions to achieve objectives (Zimmerman, 2011). In this framework, learning extends beyond knowledge acquisition to include managing one's learning journey through goal setting, progress monitoring, and strategy adjustments (Zimmerman & Martinez-Pons, 1986). SRL highlights the importance of the learner's active role in managing the learning process and developing skills to become autonomous learners. This is reflected in the comments made by Jack, Sarah, Paul, and Nancy. Jack's progress in structuring his essays shows how learning has been more than just receiving information and actively applying knowledge. Furthermore, Nancy's interest in experimenting with ChatGPT, Sarah's enthusiasm to write more and track progress, and Paul's realization and subsequent correction of his problems in writing all demonstrate the self-regulated nature of their learning.

The results are also in keeping with the principles of learner-centered instruction. Learner-centered instruction advocates a shift from a one-size-fits-all approach to education to one that emphasizes personalized and adaptive instruction, catering to the specific needs, preferences, and capabilities of learners (Richards & Schmidt, 2010). This approach is grounded in the belief that learners are unique and diverse. Therefore, to be effective, teaching methods must be customized and tailored to align with their unique learning styles, capabilities, and preferences (Griffiths & Soruç, 2021). Based on the learners' reflections, ChatGPT has successfully implemented this approach by catering to the specific needs of each learner. This is evident from the comments made by Paul and Leila, where they appreciated receiving specific feedback that catered to their unique learning problems.

6. Conclusion

The present research aimed to examine the impact of using ChatGPT 4.0 on L2 learners' motivation to write essays in English. The results revealed that, overall, interacting with ChatGPT 4.0 had a positive lasting effect on the learners' motivation to write essays in English. However, a decline in the learners' motivation at the delayed post-intervention stage, possibly stemming from an initial novelty effect, suggests the need to maintain a balance between utilizing ChatGPT as a support tool and enhancing their independent writing capabilities (Zhang, 2023). Furthermore, it indicates that although interacting with ChatGPT may enhance immediate motivation, sustained motivation requires continued support and scaffolding. To mitigate such declines, longitudinal support measures should be taken to maintain motivation following technology-mediated interventions. Finally, the learners reported an increase in motivation due to several factors, including their perceived improvement in essay writing skills, the supportive learning environment created by ChatGPT as a tutor, positive interactions with it, and the development of meta-cognitive awareness by addressing learners' specific writing issues.

It is important to approach the results of this study with caution, as there were some limitations that need to be considered. First, the data were collected from a small number of L2 learners, which could impact the applicability of the findings to a broader population. Furthermore, the subjects recruited for the study, comprising only Iranian EFL learners with B1 English proficiency,

were not sufficiently diverse. While the focus on B1 learners facilitated a controlled examination, the study's findings are primarily applicable to similar demographic groups within EFL contexts. Moreover, to represent the broad population of L2 learners of English, the sample needs to include learners with different cultural and psychological backgrounds. Additionally, the learners' individual differences were not considered in sampling the population. Taking into consideration learners' cognitive and affective aspects may add to the depth of the study's findings. Furthermore, the investigation only included self-report measures that have several limitations, such as desirability, honesty, and introspection capacity (Demetriou, Ozer & Essau, 2015). In addition, using other means of data collection, such as stimulated recall and similar measures, could have added to the diversity and richness of the findings. Another important point was that although the study involved collecting both quantitative and qualitative data at different time points in the study, to fully understand the mechanisms through which L2 learners' motivation changes, it is important to study learners over a long period of time, using longitudinal designs. Finally, our study primarily focused on L2 learners' motivation to write English essays, and we did not collect specific data on linguistic outcomes. Given the importance of studying linguistic outcomes when working with generative AI tools, including results on learners' linguistic outcomes would provide a more robust examination of the dynamics of L2 motivation and English essay writing in the face of generative AI. Hence, to investigate the impact of interacting with ChatGPT on L2 learners' motivation or other aspects of L2 development, researchers of future studies should use larger sample sizes, recruit learners with a wider range of proficiency levels, incorporate diverse L2 learning tasks, consider learners' individual differences, use other means of collecting data, adopt longitudinal designs, and include results on linguistic outcomes to enhance the trustworthiness of their analyses.

The findings offer valuable implications for SLA research and practice. They highlight the potential of incorporating AI-based tools in language classrooms to enhance L2 learners' motivation. The results may be used as a foundation for conducting comparative research on how ChatGPT and similar AI-based technologies can impact other aspects of L2 development. L2 teachers and curriculum developers may consider integrating ChatGPT or similar technologies into their L2 materials to create a supportive and adaptable learning environment that fosters motivation and leads to L2 development.

Data availability statement. Data available on request from the authors.

Authorship contribution statement. Javad Zare: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; Ahmad Al-Issa: Writing – original draft; Fatemeh Ranjbaran Madiseh: Writing – review & editing.

Funding disclosure statement. This research did not receive any specific funding.

Competing interests statement. The authors declare no competing interests.

Ethical statement. Ethical considerations, such as voluntary participation, anonymity, and confidentiality of the learners, were considered.

GenAI use disclosure statement. The authors declare no use of generative AI.

References

- Barrot, J. S. (2023) Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, (vol 57): Article 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Braun, V. & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2): 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chen, C.-M., Chen, L.-C. & Yang, S.-M. (2019) An English vocabulary learning app with self-regulated learning mechanism to improve learning performance and motivation. *Computer Assisted Language Learning*, 32(3): 237–260. <https://doi.org/10.1080/09588221.2018.1485708>
- Cohen, A. & Dörnyei, Z. (2002) Focus on the language learner: Motivation, styles and strategies. In N Schmitt, (ed.), *An introduction to applied linguistics*. London: Edward Arnold, 170–190.

- Creswell, J. W. (2014) *A concise introduction to mixed methods research*. Thousand Oaks: SAGE Publications.
- Demetriou, C., Ozer, B. U. & Essau, C. A. (2015) Self-report questionnaires. In R. L. Cautin, & S. O. Lilienfeld, (eds.), *The encyclopedia of clinical psychology*. Chichester: John Wiley & Sons, 1–6. <https://doi.org/10.1002/9781118625392.wbecp507>
- Dörnyei, Z. (2002) The motivational basis of language learning tasks. In P. Robinson, (ed.), *Individual differences and instructed language learning*. Amsterdam, John Benjamins, 137–158. <https://doi.org/10.1075/illt.2.10dor>
- Dörnyei, Z. (2005) *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah: Lawrence Erlbaum.
- Dörnyei, Z. (2019) Task motivation: What makes an L2 task engaging? In Wen, Z. & Ahmadian, M. J. (eds.), *Researching L2 task performance and pedagogy: In honor of Peter Skehan*. Amsterdam: John Benjamins, 53–66. <https://doi.org/10.1075/tblt.13.04dor>
- Dörnyei, Z., MacIntyre, P. D. & Henry, A. (eds.) (2015) *Motivational dynamics in language learning*. Bristol: Multilingual Matters.
- Dörnyei, Z. & Murphey, T. (2003) *Group dynamics in the language classroom*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511667138>
- Dörnyei, Z. & Ushioda, E. (2013) *Teaching and researching: Motivation* (2nd ed.). Abingdon: Routledge. <https://doi.org/10.4324/9781315833750>
- Fang, Z. (2021) *Demystifying academic writing: Genres, moves, skills, and strategies*. New York: Routledge. <https://doi.org/10.4324/9781003131618>
- Ferris, D. R. & Hedgcock, J. S. (2023) *Teaching L2 composition: Purpose, process, and practice* (4th ed.). New York: Routledge. <https://doi.org/10.4324/9781003004943>
- Fukuda, A. (2018). The Japanese EFL learners' self-regulated language learning and proficiency. *Journal of Pan-Pacific Association of Applied Linguistics*, 22 (1), 65–87.
- Geher, G. (2023, January 6) ChatGPT, artificial intelligence, and the future of writing. *Psychology Today*. <https://www.psychologytoday.com/au/blog/darwins-subterranean-world/202301/chatgpt-artificial-intelligence-and-the-future-of-writing>
- Griffiths, C. & Soruç, A. (2021) Individual differences in language learning and teaching: A complex/dynamic/socio-ecological/holistic view. *English Teaching & Learning*, 45(3): 339–353. <https://doi.org/10.1007/s42321-021-00085-3>
- Guo, K., Wang, J. & Chu, S. K. W. (2022) Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing*, 54: Article 100666. <https://doi.org/10.1016/j.asw.2022.100666>
- Hartwell, K. & Aull, L. (2022) Constructs of argumentative writing in assessment tools. *Assessing Writing*, 54: Article 100675. <https://doi.org/10.1016/j.asw.2022.100675>
- Hattie, J. & Timperley, H. (2007) The power of feedback. *Review of Educational Research*, 77(1): 81–112. <https://doi.org/10.3102/003465430298487>
- Jüllkunen, K. (1989) *Situation- and task-specific motivation in foreign-language learning and teaching*. Joensuu: University of Joensuu Publications in Education.
- Kormos, J. (2023) The role of cognitive factors in second language writing and writing to learn a second language. *Studies in Second Language Acquisition*, 45(3): 622–646. <https://doi.org/10.1017/S0272263122000481>
- Kormos, J. & Préfontaine, Y. (2017) Affective factors influencing fluent performance: French learners' appraisals of second language speech tasks. *Language Teaching Research*, 21(6): 699–716. <https://doi.org/10.1177/1362168816683562>
- Kormos, J. & Wilby, J. (2019) Task motivation. In Lamb, M., Csizer, K., Henry, A. & Ryan, S. (eds.), *The Palgrave handbook of motivation for language learning*. Cham: Palgrave Macmillan, 267–286. https://doi.org/10.1007/978-3-030-28380-3_13
- Lamb, M. (2016) Motivation. In Hall, G. (ed.), *The Routledge handbook of English language teaching*. London: Routledge, 324–338. <https://doi.org/10.4324/9781315676203-28>
- Lincoln, Y. S. & Guba, E. G. (1985) *Naturalistic inquiry*. Beverly Hills: Sage Publications.
- Ma, J. H. (2009) *Autonomy, competence, and relatedness in L2 learners' task motivation: A self-determination theory perspective*. University of Hawai'i at Manoa, unpublished doctoral dissertation.
- McMurtrie, B. (2022, December 13) AI and the future of undergraduate writing. *The Chronicle of Higher Education*. https://www.chronicle.com/article/ai-and-the-future-of-undergraduate-writing?cid=gen_sign_in
- McNeill, K. L. & Krajcik, J. (2009). Synergy between teacher practices and curricular scaffolds to support students in using domain-specific and domain-general knowledge in writing arguments to explain phenomena. *The Journal of the Learning Sciences*, 18(3), 416–460. <https://doi.org/10.1080/10508400903013488>
- Mozgalina, A. (2015) More or less choice? The influence of choice on task motivation and task engagement. *System*, 49: 120–132. <https://doi.org/10.1016/j.system.2015.01.004>
- Olsen, T. & Hunnes, J. (2024) Improving students' learning—The role of formative feedback: Experiences from a crash course for business students in academic writing. *Assessment & Evaluation in Higher Education*, 49(2): 129–141. <https://doi.org/10.1080/02602938.2023.2187744>
- Plonsky, L. & Oswald, F. L. (2014) How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4): 878–912. <https://doi.org/10.1111/lang.12079>

- Poupore, G. (2013) Task motivation in process: A complex systems perspective. *The Canadian Modern Language Review*, 69(1): 91–116. <https://doi.org/10.3138/cmlr.1139>
- Richards, J. C. & Schmidt, R. W. (2010) *Longman dictionary of language teaching and applied linguistics* (4th ed.). Harlow: Longman.
- Rudolph, J., Tan, S. & Tan, S. (2023) ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1): 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Ryan, R. M. & Deci, E. L. (2017) *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: The Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- Seegers, G. & Boekaerts, M. (1993) Task motivation and mathematics achievement in actual task situations. *Learning and Instruction*, 3(2): 133–150. [https://doi.org/10.1016/0959-4752\(93\)90012-O](https://doi.org/10.1016/0959-4752(93)90012-O)
- Skehan, P. (1989) *Individual differences in second-language learning*. London: Edward Arnold.
- Su, Y., Lin, Y. & Lai, C. (2023) Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57: Article 100752. <https://doi.org/10.1016/j.asw.2023.100752>
- Ushioda, E. (2008) Motivation and good language learners. In Griffiths, C. (ed.), *Lessons from good language learners*. Cambridge: Cambridge University Press, 19–34. <https://doi.org/10.1017/CBO9780511497667.004>
- Wang, Y., Derakhshan, A., Pan, Z. & Ghiasvand, F. (2023) Chinese EFL teachers' writing assessment feedback literacy: A scale development and validation study. *Assessing Writing*, 56: Article 100726. <https://doi.org/10.1016/j.asw.2023.100726>
- Weaver, M. R. (2006) Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education*, 31(3): 379–394. <https://doi.org/10.1080/02602930500353061>
- Wei, X., Zhang, L. J. & Zhang, W. X. (2020). Associations of L1-to-L2 rhetorical transfer with L2 writers' perception of L2 writing difficulty and L2 writing proficiency. *Journal of English for Academic Purposes*, 47, 100907. <https://doi.org/10.1016/j.jeap.2020.100907>
- Weinstein, C. E., Acee, T. W. & Jung, J. (2011) Self-regulation and learning strategies. *New Directions for Teaching and Learning*, 2011(126): 45–53. <https://doi.org/10.1002/tl.443>
- Zare, J. & Aqajani Delavar, K. (2024) Enhancing English learning materials with data-driven learning: A mixed-methods study of task motivation. *Journal of Multilingual and Multicultural Development*, 45(9): 4011–4027. <https://doi.org/10.1080/01434632.2022.2134881>
- Zhang, R. (2023) 生成式人工智能技术对教育领域的影响——关于ChatGPT的专访 [The impact of generative artificial intelligence technology on the education field: An interview with ChatGPT]. *Journal of Educational Research*, 44(2): 5–14. <https://doi.org/10.13811/j.cnki.eer.2023.02.001>
- Zhang, Z. & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon*, 10(3). <http://dx.doi.org/10.1016/j.heliyon.2024.e25370>
- Zimmerman, B. J. (2008) Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1): 166–183. <https://doi.org/10.3102/0002831207312909>
- Zimmerman, B. J. (2011). Motivational sources and outcomes of self-regulated learning and performance. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 49–64). Routledge.
- Zimmerman, B. J. & Martinez-Pons, M. (1986) Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23(4): 614–628. <https://doi.org/10.3102/00028312023004614>

About the authors

Javad Zare is an associate professor of applied linguistics at Kosar University of Bojnord, Bojnord, Iran. His most recent publications have appeared in *Applied Linguistics*, *Brain Sciences*, *British Journal of Educational Studies*, *Computer Assisted Language Learning*, *Contemporary Educational Psychology*, *Current Psychology*, *Discourse Processes*, *Educational Review*, *English for Specific Purposes*, *Frontiers in Psychology*, *Innovation in Language Learning and Teaching*, *International Journal of Applied Linguistics*, *Journal of English for Academic Purposes*, *Journal of Pragmatics*, *Journal of Second Language Writing*, *Language Teaching Research*, *Lingua*, *ReCALL*, *System*, and *Text & Talk*. His research areas of interest include positive psychology, computer-assisted language learning, task-based language teaching, corpus linguistics, and English for academic purposes.

Ahmad Al-Issa is a professor of English linguistics at the American University of Sharjah in the United Arab Emirates. His research interests are in the areas of sociolinguistics, intercultural communication, language and identity, global English, pragmatics, language planning, and language policy.

Fatemeh Ranjbaran Madiseh is an assistant lecturer at Sultan Qaboos University, Muscat, Oman. Her research areas of interest include but are not limited to educational technology, computer-assisted language learning, English for specific purposes, and cognitive diagnostic assessment.