Contemporary Educational Technology, 2025, **17**(1), ep548
ISSN: 1309-517X (Online)

OPEN ACCESS

Research Article

Leveraging AI to enhance writing skills of senior TFL students in Kazakhstan: A case study using "Write & Improve"

Abbas Bodaubekov [1]
0009-0001-7959-2605

Shakhrizat Agaidarova [1]
0000-0003-0047-9886

Talgat Zhussipbek [2,3]
0000-0002-6701-4701

Davronzhon Gaipov [1]
0000-0003-3012-8067

Nuri Balta [1*]
0000-0002-6364-5346

[1] SDU University, Almaty, KAZAKHSTAN
[2] Karaganda Buketov University, Karaganda, KAZAKHSTAN
[3] Michigan State University, East Lansing, MI, USA
[*] Corresponding author: baltanuri@gmail.com

Citation: Bodaubekov, A., Agaidarova, S., Zhussipbek, T., Gaipov, D., & Balta, N. (2025). Leveraging AI to enhance writing skills of senior TFL students in Kazakhstan: A case study using "Write & Improve". *Contemporary Educational Technology, 17*(1), ep548. https://doi.org/10.30935/cedtech/15687

ARTICLE INFO

Received: 11 Aug 2024

Accepted: 6 Nov 2024

ABSTRACT

This study investigates the effectiveness of feedback provided by teachers versus feedback generated by the Write & Improve platform in enhancing the writing skills of senior undergraduate students enrolled in a "two foreign language" program at a private university in Kazakhstan. The quasi-experimental design involved four teachers, each teaching one control and one experimental class, totaling eight groups of students. Pre- and post-tests were conducted over a period of five weeks, focusing on task achievement, coherence and cohesion, lexical resource, grammar and accuracy, and overall score. Data analysis included descriptive statistics, Mann-Whitney U tests for pre-test comparisons, and MANCOVA analyses for post-test comparisons. Results show no significant difference in the impact of Write & Improve feedback compared to traditional teacher feedback across multiple dimensions of the writing test, both within individual teachers' classes and when combined. Longitudinal analysis reveals fluctuating scores over time with no consistent improvement. Thus, the study concludes that the Write & Improve tool is equally effective as teacher feedback in improving students' writing skills. This implies that educational institutions can potentially integrate technology-based feedback systems like Write & Improve alongside traditional teaching methods to enhance student learning outcomes in writing proficiency.

Keywords: AI feedback, English language, student learning outcomes, teacher feedback, Write & Improve tool

INTRODUCTION

The application of artificial intelligence (AI) in education has the potential to transform education by offering personalized, adaptable, and efficient educational materials and assistance to students of all levels (Relmasira et al. 2021). Educators may better identify students' specific requirements, personalize

instructional content, and improve overall educational achievements by utilizing AI technologies. Numerous scholars investigated the use of AI tools in teaching and learning in relation to various platform use (de los Ángeles Domínguez-González et al., 2023; Fitria, 2021), digital pedagogy (Timonen & Ruokamo 2021; Väätäinen & Ruokamo, 2021), digital storytelling (Niemi & Niu, 2021; Robin, 2008), advantages of using AI in education (Luckin & Holmes, 2016), and advantages of AI writing tools (Toscano, 2023). For example, Luckin and Holmes (2016) argue that integrating AI tools in education enhances inclusivity by assisting disabled students, offering immediate and detailed feedback, analyzing a broader audience's work, and providing personalized feedback based on individual strengths and weaknesses. Utilizing such AI tools develops more student-centered learning, thus enhancing the overall academic experience of students.

On the other hand, traditional teaching methods, also known as teacher-centered instruction, are a pedagogical technique in which the instructor plays a pivotal role in transferring information, guiding students through planned lessons, and measuring their progress by giving marks. Recent studies have shown many disadvantages of applying such traditional teaching methods in education, claiming them to be outdated, teacher-centric, monopolized, and limiting students' self-esteem and motivation. Moreover, within the traditional way of teaching students become passive listeners learning from observing and listening to the teacher, who is in the center of the class (Hadžimehmedagić & Akbarov, 2014; Kompa, 2012). In contradiction, Schwerdt and Wuppermann (2011) argue that traditional teaching methods are not uniformly less effective than modern alternative methods. They conducted a study comparing the academic performance of the same students when exposed to both traditional and non-traditional teaching approaches. Their findings reveal that, in certain cases, traditional teaching was equally or even more effective than modern approaches. This challenges the oversimplified notion of teaching effectiveness and suggests that traditional methods may have value in specific contexts.

As was previously stated, using AI tools in the classroom improves students' academic achievement by scaffolding them in a variety of ways, including by giving meticulous feedback, personalizing the educational materials they receive, responding quickly, and identifying their unique needs and areas for improvement (Hervás Gómez & Toledo Morales, 2018). This study investigates the impact of a such AI-assisted writing tool Write & Improve on the development of students' writing skills within Kazakhstani context. According to Karpova (2020), the AI tool Write & Improve provides feedback on students' writing work based on the common European framework of reference for languages, employing four evaluation criteria to assess the students' work. The criteria include content, communicative achievement, organization and language. Wali and Huijser (2018) provide a detailed review of the AI tool Write & Improve claiming it to be user-friendly, flexible, accessible and customizable.

The use of AI in education has gained momentum, particularly in the field of teaching English as a foreign language, where a multitude of recent studies have investigated the impact of the AI-driven tool Write & Improve on enhancing students' writing skills, providing valuable suggestions for further implementation (Cheng, 2017; Curry & Riordan, 2021; Heydari & Marefat, 2020; Karpova, 2020; Taskiran & Goksel, 2022; Toscano, 2023; Wang et al., 2013; Wali & Huijser, 2018). For instance, Curry and Riordan (2021) and Heydari and Marefat (2020) investigated the provision of automated feedback on students' writing. Both studies underscore the effectiveness of using Write & Improve in improving students' writing proficiency. It is interesting to note that, both studies highlight the benefit of using AI-driven tool in developing basic mistakes in spelling and grammar (Curry & Riordan, 2021) and enhancing writing at the word level instead of focusing on the sentence level (Heydari & Marefat, 2020). Another series of studies conducted in relation to the use of Write & Improve, Karpova (2020) and Wali and Huijser (2018) claim AI-driven tool to be profoundly beneficial not only to the development of writing skills, but also to fostering students' motivation in writing. Additionally, Cheng (2017), Taskiran and Goksel (2022), and Wang et al. (2013) also investigated the application of AI-assisted automated writing evaluation (AWE) and online automated feedback systems on its effectiveness in using it to develop students' writing skills in English as a foreign language course. For instance, Cheng's (2017) study found that participants in the experimental group significantly outperformed those in the control group in terms of their overall scores (OS) in writing reflective journals. Similarly, Wang et al. (2013) discovered that AWE system developed students' writing accuracy, autonomy and student interaction while learning English as a foreign language. On the other hand, Taskiran and Goksel (2022) examined effectiveness and success in students' writing based on the feedback obtained both from the teacher and from the automated feedback

tool. The results showed that participants improved their academic writing skills through regular feedback from the teacher, these results are marginally better than using automated software for the feedback. Conversely, Toscano's (2023) research findings align with those of earlier ones done by Curry and Riordan (2021), Heydari and Marefat (2020), Cheng (2017) and Wang et al. (2013). Toscano (2023) found that students significantly improved their average score in four assessment criteria, namely content, communicative achievement, organization and language. An AI-driven tool Write & Improve dramatically improved students' writing performance and had positive effect on students' writing skills (Toscano, 2023).

Understanding students' perceptions of using AI-driven tools in learning is crucial for shaping quality education (Yilmaz et al., 2023). It helps in assessing the effectiveness and identifying potential challenges related to the integration of AI tools in education (Almaiah et al., 2022; Alammari & Abdel-Reheem Amin, 2023; Chan & Hu, 2023; Hussain, 2020; Sit et al., 2020). A growing awareness of the potential benefits of integrating AI-driven tools in education has been reported by recent studies. According to Chan and Hu (2023), Hussain (2020), and Alammari and Abdel-Reheem Amin (2023), the integration of AI-driven tools positively impacts students' effective learning outcomes. For instance, Chan and Hu (2023) found positive students' attitude towards AI as it provides personalized learning support and writing assistance. The AI tool has been used as a brainstorming and data analysis tool which scaffolded students' education in the context of higher education. Moreover, Hussain (2020) also highlights favorable perceptions of the students towards AI tools and add its time saving advantage making the learning and teaching process more productive and efficient. Almost 75% of the research participants underscore the potential benefits of using AI as providing meaningful feedback, flexible learning, and unbiased assessment. Though students expressed an abundance of positive perceptions towards AI use in learning, some stated certain concerns such as privacy, reliability, plagiarism, and data security were among them (Chan & Hu, 2023).

Another series of studies also focused on students' perceptions of AI-driven tool's usage in education in relation to various contents (Alammari & Abdel-Reheem Amin, 2023; Sit et al., 2020) and students' anxiety (Almaiah et al., 2022). The use of AI-driven technologies in healthcare provision (Sit et al., 2020) and English as a foreign language (Alammari & Abdel-Reheem Amin, 2023) is perceived positively by students, indicating favorable learning results. Alammari and Abdel-Reheem Amin (2023) assert that EFL students' attitudes about adopting AI technologies were generally favorable. Students' academic writing abilities improve for the better. Namely, AI has been applied to paraphrase, enhance sentence structure, rectify grammar, and discover synonyms. The effectiveness of employing AI tools to enhance writing organization, raise writing confidence, expand vocabulary, improve grammar, and increase source knowledge is widely agreed upon by research participants. Sit et al. (2020) reported similar positive impressions when examining medical students' perspectives towards the use of AI in radiography. Nearly all research participants (88%) believed that AI had a significant role in healthcare. Even if the positive findings exceed the negative ones, some medical students were discouraged from pursuing careers in radiology because they were concerned that AI might eventually replace them. This was a fear shared by research participants conducted by Sit et al. (2020). Additionally, word choice errors, contextual misunderstandings, and an excessive dependence on the technology were concerns raised by EFL students in the study conducted by Alammari and Abdel-Reheem Amin (2023). Another issue expressed by research participants about AI technologies is anxiety, according to Almaiah et al. (2022). The study investigated how students' desire to use AI technologies in education is affected by anxiety. The findings indicated that students who experience computer anxiety had worse self-efficacy when utilizing AI technology. According to Almaiah et al. (2022), students are more interested in utilizing AI technologies in education when they are less apprehensive and when aspects such as immersion, motivation, interaction, and satisfaction are changed positively. Almaiah et al. (2022) recommend that phycological concerns be considered while implementing AI technology in education since it may impede the learning process. Overall, the studies mentioned above demonstrate that integrating AI tools into the educational environment can cultivate students' learning outcomes in a number of ways, including improving writing abilities, increasing vocabulary, boosting confidence, and encouraging productive learning. However, there are some potential drawbacks, including anxiety, contextual misunderstanding, and an excessive dependence on technology.

Research indicates a wealth of studies focusing on teachers' experience of using AI-driven tools in the educational domain. Some researchers have delved into teachers' attitudes towards the incorporation of AI in English language teaching (Chounta et al., 2022; Sumakul et al., 2022), while others have scrutinized the

application of AI-based tools for evaluation purposes (Luckin, 2017). Moreover, some empirical studies have concentrated on teachers' conceptualizations of AI's role in education (Yau et al., 2023), whereas others have measured teachers' trust in using AI technologies (Nazaretsky et al., 2022). Additionally, a separate body of works have explored the multitude of factors that affect teachers' perspectives in this context (Chounta et al., 2022; McGrath et al., 2023).

Regarding teachers' perspective on the use of AI-powered tools in education, the research results are in consistent with studies that focus on students' perceptions of incorporating AI-driven tools in learning, and the overall results are mostly positive (Chounta et al., 2022; Sumakul et al., 2022). For example, Sumakul et al. (2022) looked at the pros and cons of using AI in EFL instruction. Overall, the study found that most participants had favorable opinions towards AI in EFL classes. Teachers saw AI favorably since it supported the development of grammar, pronunciation, and vocabulary. In a similar vein, Chounta et al. (2022) discovered that the views of AI use in Estonian schools were generally favorable. The teachers emphasized the benefits of utilizing AI, such as satisfying a variety of student demands, helping with administrative tasks, saving time, and providing tailored support for each student. Despite the numerous advantages of using AI in education, the studies emphasize certain challenges and concerns regarding its implementation in classrooms. Challenges include the possibility of an excessive dependence on AI, ethical issues, and the necessity of human interaction in EFL were noted by Sumakul et al. (2022). Furthermore, Chounta et al. (2022) have recognized many problems, such as data security, privacy policies, and the possibility of AI taking the place of instructors.

Another study conducted by Yau et al. (2023) investigated how teachers conceptualize AI education using a phenomenon-graphic approach. The study investigated how teachers perceive AI education and classified their answers into six categories. Those six categories of conceptualization include "technology bridging, knowledge delivery, interest stimulation, ethics establishment, capability cultivation, and intellectual development" (Yau et al., 2023, p. 1041). The relationship between these categories is hierarchical, going from a basic surface-level understanding to a more detailed and comprehensive understanding (Yau et al. 2023). Interestingly, plenty of study results indicate a need for teacher training in the field of AI implementation in the classroom. Yau et al. (2023) advocate that policymakers address the needs of non-technical teachers in order to improve their readiness to educate with AI. According to Sumakul et al. (2022), instructors require training programs in order to effectively use AI tools in their teaching methods. Similarly, Chounta et al. (2022) emphasize the benefits of professional development programs for using AI technologies, particularly for instructors who lack essential IT skills.

As previously said, technological competency is one of the primary factors impacting teachers' opinions about the implementation of AI tools in education, according to research by Chounta et al. (2022) and Yau et al. (2023). The more teachers are tech savvy the more they hold positive perceptions towards AI usage in teaching. Moreover, Chounta et al. (2022) spotted that teachers' age, teaching experience and their technical competence as stated above are some of the crucial factors that affect their perception. Other factors that affect teachers' perceptions relate to students type such as whether AI implementation will be with first-generation students, or other age range, or students with special educational needs (McGrath et al., 2023). McGrath et al.'s (2023) research also highlight the importance of instructors' AI expertise as a critical aspect in promoting the effective use of AI in the teaching and learning processes. Overall, factors such as teachers' age, teaching experience, technical competence, and student profiles are major factors that influence teachers' perceptions of implementing AI tools in education provision.

In an AI-driven education, teachers' knowledge of AI and their trust in AI tools and using AI in assessment play a significant role in shaping quality education. Comprehension of these intertwined concepts is critical to the potential development of AI-driven education. Luckin (2017) state that AI tools are beneficial in terms of their usage in personalized assessment. According to Luckin (2017), AI can create evaluating instruments based on students' individual needs and provide instant feedback as was the case in the study conducted by Karpova (2020) in relation to AI tool Write & Improve. Moreover, Luckin (2017) claims that AI assessing tools can replace formal, traditional high-stake exams with more formative and continuous assessment. A separate study accomplished by Nazaretsky et al. (2022) developed an instrument that measures teachers' trust in AI-based educational tool. The scholars claim that their developed instrument is reliable in measuring teachers' trust in AI tools. The study results reveal eight key factors that affect teachers trust in using AI tools in teaching.

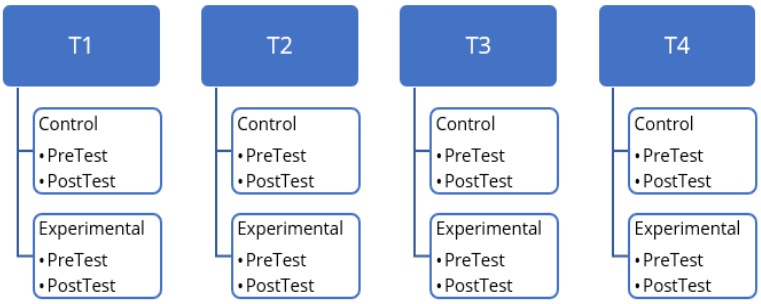

Figure 1. Study design (Source: Authors)

Overall, integrating AI tools in teaching for enhancing educational outcomes of the students is beneficial in terms of providing instant feedback, personalized support, and time saving opportunities though some challenges exist.

The objective of this study was to contrast two distinct types of feedback. The first type involved feedback from teachers, while the second type involved feedback from an AI tool. Our study was guided by the following research question: What is the impact of integrating the AI writing tool Write & Improve on the writing skills of 1st year teaching foreign language (TFL) students in Kazakhstan compared to traditional assessment methods?

METHODOLOGY

The selection of the methodology was driven by the desire to create a controlled environment that reflects actual teaching scenarios while still maintaining the ability to assess the impact of different feedback mechanisms. By using a quasi-experimental design, we were able to accommodate practical constraints such as class sizes and instructor preferences, while still obtaining valuable comparative data (Shadish et al., 2002). Furthermore, this design was selected to ensure that the feedback mechanisms were applied in a manner consistent with everyday classroom practices, thereby enhancing the ecological validity of the study's findings.

Study Design

This quasi-experimental design was carried out with four teachers each having two classes of one control and one experimental. In other words, there were eight groups of students, who were pre and posted tested, participated in this study. The design of the study is depicted in **Figure 1**.

Participants and Context

The participants were 134 1st year TFL students enrolled in the "two foreign language" program at a private university in Kazakhstan, comprising 17 males and 117 females. They ranged in age from 18 to 21 years old and were ethnically diverse, including Kazakh, Uzbek, Russian, and Uighur students, most of whom were born in Kazakhstan, with some international students from Tajikistan and Turkey. The majority of students are trilingual, proficient in Kazakh and Russian, while international students are proficient in their native language and Kazakh or Russian. Before enrolling in the university, they had studied English as a foreign language for approximately 7–10 years. In Kazakhstan, English is taught from the 5th grade up to the 12th grade in secondary school. At the university, 90% of the courses in their educational programs were conducted in English. The "two foreign language" program trains teachers of English as a foreign language as a major, with an additional foreign language as a minor. The participants' language proficiency levels ranged from intermediate to upper-intermediate.

The study course, English for specific purposes, was taught by four instructors who hold master's degrees in TESOL. The course aimed to help students acquire language proficiency tailored to diverse contexts, facilitating familiarity with conversational English in daily life and various cultural settings. The curriculum was formulated to enhance students' proficiency in communicative language skills up to an advanced level, incorporating all four language domains: reading, speaking, listening, and writing. The course comprised three-hour classes per week, lasting 15 weeks from the beginning of September to the middle of December.

The experiment was conducted from the 8th week, starting at the end of October and ending in the 13th week, at the end of November. Participants' classes were randomly divided into either the control group (65) or the experimental group (69).

Since participants in the experimental group had no prior experience using the Write & Improve platform before the intervention, the researcher provided a short training on how to use it.

Instrument, Implementation, and Data Collection

The Write & Improve tool, an AI-assisted platform, was utilized to provide feedback on students' writing. This tool offered feedback based on various assessment criteria, including content, coherence, lexical resource (LR), and grammar and accuracy (GA), allowing for a comparison between AI-generated feedback and traditional teacher feedback.

At the beginning of the study, the researcher discussed with the instructors the specifics of the experiment. Written assignments, according to their experience, were not the most desirable type of students' assignments. In order to motivate participants to actively take part, the instructors agreed to provide bonus points for participation and completion of written assignments. Moreover, in order to reduce stress for participants due to the academic workload, except for the pre and post-tests, it was decided to allow complete tasks (one task per week) at a convenient time for participants. A deadline (the day before they have a class) for the submission of written papers was set for both control and experimental groups.

After discussions and agreements with the instructors, the participants were randomly divided into control and experimental groups. The researcher visited each group and explained the project, informing them that the research concerns the quality of written work assessment. Different assessment methods would be employed, and eventually, a comparison would be made between the results. Specifically, the experimental group was informed that they would work with the Write & Improve platform. The researcher, along with the instructors, demonstrated how the platform functions and its operational principles.

Additionally, participants were instructed to independently perform the tasks and to refrain from any form of plagiarism. To ensure originality, works were checked using the Grammarly anti-plagiarism tool. Participants were also informed that tasks must be completed at any time before the deadline, with a writing time limit of 20 minutes. This duration corresponds to the time allocated for task 1 type assignments in the IELTS writing test.

In the first week, a pre-test was administered to each group according to the course schedule. Similar to all other writing assignments, the pre-test involved a writing task resembling the IELTS writing task 1. All tasks were sourced from the Write & Improve platform, specifically from the IELTS academic section.

Five writing tasks were utilized in the present study for both the experimental and control groups. These tasks were selected from the collection of assignments available on the Write & Improve platform in the IELTS Academic section. The writing tasks entailed composing a 150-word essay describing a provided chart or table, akin to the IELTS writing task 1.

The writing assignments were evaluated according to the International English Language Testing System (IELTS) writing task assessment criteria (https://ielts.org/organisations/ielts-for-organisations/understanding-ielts-scoring). The effectiveness and reliability of IELTS writing assessment criteria are generally established through rigorous testing, validation studies, and ongoing quality assurance measures. The IELTS test undergoes continuous research and development to ensure that it remains a reliable and valid measure of a test taker's English language proficiency. Each IELTS test includes four parts: task achievement (TA), coherence and cohesion (CC), LR, and GA.

During the experiment, except for pre and post-test, all communication with researcher was held online. For control group Google Classroom was chosen as a platform for work submission. With experimental group feedback was not given by instructors but only by Write & Improve platform. The post-test was performed again classes with each group in turn, depending on the schedule of the course. All experiments took five weeks with one week for one writing assignment.

Data Analysis

Initially, descriptive statistics, including measures such as mean, standard deviation, and normality, were provided. Subsequently, the pre-test scores of both the control and experimental groups were compared using the Mann-Whitney U test, considering the non-normal distribution of the data. To ensure the initial equality of the groups, the Mann-Whitney U test was employed to compare students' pre-test scores in the control and experimental groups. Following that, MANCOVA analyses were conducted to assess and compare the post-test scores of the control and experimental groups. Lastly, graphical representations, including trend lines, were utilized to illustrate the change in students' scores over time.

Ethics

Ethical considerations are central to this study, ensuring that participant rights, privacy, and well-being are protected throughout the research process. Ethical approval was obtained from the SDU University Research Committee, which reviews studies classified as "no more than minimal risk research." Participants were informed about the study's purpose, procedures, and any potential risks, allowing for fully informed consent. Measures to maintain anonymity and confidentiality were implemented, including data coding, secure storage, and access restriction to authorized personnel only. Data privacy is further safeguarded through encryption, ensuring compliance with institutional and ethical standards. Additionally, the study balanced AI-driven feedback with traditional instructional methods to prevent overreliance on technology, fostering a supportive educational environment. Regular monitoring and open channels for participant feedback were maintained to address any concerns promptly, upholding the integrity and ethical standards of the research.

RESULTS

The study evaluated the impact of feedback from the Write & Improve tool in comparison to feedback from teachers in classrooms led by four coded teachers: T1, T2, T3, and T4. The analysis of Write & Improve feedback, as opposed to traditional teacher feedback, will be presented for each teacher using a one-way MANCOVA. Each teacher oversaw two classes, one receiving control (traditional teacher feedback) and the other experimental (Write & Improve feedback).

Descriptive Statistics

To assess the impact of the Write & Improve tool, pre-tests and post-tests covering four dimensions (TA, CC, LR, GA, and OS) were conducted. **Table 1** displays the number of students (N), mean (M), standard deviation (SD), normality (p), and group type (CG: control group and EG: experimental group) for each teacher.

Table 1. Descriptive statistics of groups

	Score	Group	N	M	SD	p		Score	Group	N	M	SD	p
T1-Pre	TA	EG	17	3.59	0.54	0.06	T1-Post	TA	EG	14	3.86	0.5	0.09
		CG	9	4.72	0.44	< .001			CG	9	3.67	0.35	0.02
	CC	EG	17	3.53	0.41	0.01		CC	EG	14	3.79	0.58	0.12
		CG	9	3.78	0.36	0.01			CG	9	3.67	0.50	0.36
	LR	EG	17	3.5	0.47	0.01		LR	EG	14	4.04	0.66	0.09
		CG	9	3.61	0.49	0.00			CG	9	3.78	0.51	0.05
	GA	EG	17	3.56	0.53	0.00		GA	EG	14	3.64	0.53	0.06
		CG	9	3.00	0	NA			CG	9	4.00	0.50	0.08
	OS	EG	16	3.52	0.32	0.10		OS	EG	14	3.83	0.44	0.87
		CG	9	3.78	0.2	0.03			CG	9	3.78	0.38	0.15
T2-Pre	TA	EG	19	3.34	0.8	0.01	T2-Post	TA	EG	12	4.42	0.47	0.00
		CG	19	4.05	0.69	0.01			CG	21	4.02	0.68	0.00
	CC	EG	19	3.66	0.55	0.01		CC	EG	12	4.21	0.45	0.13
		CG	19	4.11	0.57	0.01			CG	21	3.71	0.75	0.30
	LR	EG	19	3.87	0.6	0.01		LR	EG	12	4.58	0.47	0.00
		CG	19	4.05	0.55	0.00			CG	21	4.17	0.76	0.00
	GA	EG	19	3.92	0.48	0.01		GA	EG	12	4.21	0.26	< .001
		CG	19	3.32	0.48	< .001			CG	21	3.98	0.66	< .001
	OS	EG	19	3.70	0.49	0.20		OS	EG	12	4.35	0.29	0.40
		CG	19	3.88	0.42	0.63			CG	21	3.97	0.66	0.01

Table 1. (continued)

	Score	Group	N	M	SD	p		Score	Group	N	M	SD	p
T3-Pre	TA	EG	14	3.50	0.78	0.01	T3-Post	TA	EG	8	4.44	0.62	0.02
		CG	17	3.41	0.85	0.06			CG	17	3.76	0.59	< .001
	CC	EG	14	3.50	0.55	0.00		CC	EG	8	4.31	0.37	0.03
		CG	17	3.50	0.61	0.06			CG	17	3.88	0.78	0.28
	LR	EG	14	3.75	0.67	0.11		LR	EG	8	4.56	0.50	0.01
		CG	17	3.56	0.68	0.00			CG	17	4.09	0.78	0.00
	GA	EG	14	3.71	0.43	0.07		GA	EG	8	4.31	0.53	0.37
		CG	17	3.26	0.5	< .001			CG	17	3.94	0.86	0.14
	OS	EG	14	3.62	0.43	0.82		OS	EG	8	4.41	0.43	0.34
		CG	17	3.43	0.51	0.09			CG	17	3.92	0.69	0.01
T4-Pre	TA	EG	21	3.55	0.96	< .001	T4-Post	TA	EG	14	4.00	0.44	0.02
		CG	16	3.09	0.76	0.08			CG	12	3.71	0.50	0.03
	CC	EG	21	3.24	0.85	0.00		CC	EG	14	3.75	0.55	0.09
		CG	16	3.31	0.6	0.04			CG	12	3.58	0.42	0.01
	LR	EG	21	3.60	0.74	< .001		LR	EG	14	4.11	0.49	< .001
		CG	16	3.13	0.43	0.00			CG	12	3.92	0.70	0.11
	GA	EG	21	3.40	0.82	0.00		GA	EG	14	3.86	0.53	0.01
		CG	16	3.47	0.64	< .001			CG	12	3.67	0.58	0.07
	OS	EG	21	3.45	0.77	< .001		OS	EG	14	3.93	0.42	0.43
		CG	16	3.25	0.41	0.10			CG	12	3.72	0.48	0.18

Table 2. T1–Pre-/post-test comparisons

	Statistic	p
Task achievement	9.5	< .001
Coherence and cohesion	50.5	0.131
Lexical resource	65	0.532
Grammar and accuracy	27	0.005

Table 3. Multivariate tests for T1

	Test	Value	F	df1	df2	p
T1-Post	Pillai's trace	0.29	1.60	4	16	0.222
	Wilks' lambda	0.71	1.60	4	16	0.222
Task achievement	Pillai's trace	0.37	2.34	4	16	0.099
	Wilks' lambda	0.63	2.34	4	16	0.099
Grammar and accuracy	Pillai's trace	0.35	2.11	4	16	0.127
	Wilks' lambda	0.65	2.11	4	16	0.127

Inferential Statistics

For each teacher, we initiated the examination of the contrast between pre-test scores of control and experimental groups using the Mann Whitney U test due to the non-normal distribution of data. This approach was adopted for two primary reasons. Firstly, it is essential to have knowledge of the pre-test scores for both control and experimental groups when comparing their post-test scores, as this helps determine the initial success levels of the groups. Secondly, in instances of disparities, identifying covariates for MANCOVA analyses becomes crucial.

Teacher 1

Table 2 displays the pre-test scores of both experimental and control groups across the dimensions of the test administered to students following traditional feedback and Write & Improve feedback. Due to the absence of normality, as indicated in **Table 1**, the Mann-Whitney U test was utilized.

As seen in **Table 2**, control and experimental groups are not initially equal on TA ($p < .001$), and GA ($p = .005$). These two dimensions are used as covariate in the following MANCOVA analyses (**Table 3**).

Table 3 discloses that, regardless of whether covariates are considered or not, we found no statistically significant overall difference in the means across multiple dependent variables simultaneously ($p > 0.05$). In simpler terms, there is no distinction between the impact of Write & Improve tool feedback and traditional teacher feedback on students' scores across various dimensions of the test in the classes of the first teacher.

Table 4. T2–Pre-/post-test comparisons

	Statistic	p
Task achievement	96	0.010
Coherence and cohesion	107.5	0.026
Lexical resource	149.5	0.319
Grammar and accuracy	73	< .001

Table 5. Multivariate tests for T2

	Test	Value	F	df1	df2	p
T2-Post (EG=1 & CG=2)	Pillai's trace	0.15	1.16	4	26	0.353
	Wilks' lambda	0.85	1.16	4	26	0.353
Task achievement (3)	Pillai's trace	0.20	1.63	4	26	0.196
	Wilks' lambda	0.80	1.63	4	26	0.196
Grammar and accuracy (3)	Pillai's trace	0.10	0.70	4	26	0.599
	Wilks' lambda	0.90	0.70	4	26	0.599

Table 6. T3–Pre-/post-test comparisons

	Statistic	p
Task achievement	105	0.574
Coherence and cohesion	119	1.000
Lexical resource	94.5	0.318
Grammar and accuracy	56.5	0.008

Table 7. Multivariate tests for T3

	Test	Value	F	df1	df2	p
T3-Post (EG=1 & CG=2)	Pillai's trace	0.30	2.08	4	19	0.124
	Wilks' lambda	0.70	2.08	4	19	0.124
Grammar and accuracy (3)	Pillai's trace	0.36	2.70	4	19	0.062
	Wilks' lambda	0.64	2.70	4	19	0.062

Teacher 2

Table 4 presents the pre-test scores of experimental and control groups on the dimensions of the test administered to students after the traditional feedback and Write & Improve feedback. Since there was no normality, referring to **Table 1**, Mann-Whitney U Test was employed.

As seen in **Table 4**, control and experimental groups are not initially equal on TA ($p = .01$), CC ($p < .001$), and GA ($p < .001$). Only TA, and GA are used as covariate in the following MANCOVA analyses (**Table 5**) because TA, and CC had significant correlation between them ($r = 0.68$).

Table 5 reveals that, both with and without the inclusion of covariates, we observed no statistically significant overall difference in means across multiple dependent variables simultaneously ($p > .05$). Essentially, this implies that there is no discernible impact of the Write & Improve tool's feedback, in comparison to traditional feedback following the assessment, on students' scores across the various dimensions of the test in the classes of the second teacher.

Teacher 3

Table 6 presents the pre-test scores of experimental and control groups on the dimensions of the test administered to students after the traditional feedback and Write & Improve feedback. Since there was no normality, referring to **Table 1**, Mann-Whitney U test was employed.

As seen in **Table 6**, control and experimental groups are not initially equal to GA ($p = .008$). Thus, it is used as a covariate in the following MANCOVA analyses (**Table 7**).

Table 7 indicates that with and without the covariate, we did not find a significant overall difference in the means across multiple dependent variables simultaneously ($p > .05$). In other words, there is no effect of Write & Improve tool's feedback, when compared to traditional feedback after the assessment, on students' scores on the dimensions of the test in the classes of the third teacher.

Table 8. T4–Pre-/post-test comparisons

	Statistic	p
Task achievement	92.5	0.017
Coherence and cohesion	164.5	0.924
Lexical resource	69.5	0.002
Grammar and accuracy	167.5	1.000

Table 9. Multivariate tests for T4

	Test	Value	F	df1	df2	p
T4-Post (EG=1 & CG=2)	Pillai's trace	0.14	0.84	4	20	0.517
	Wilks' lambda	0.86	0.84	4	20	0.517
Task achievement	Pillai's trace	0.47	4.52	4	20	0.009
	Wilks' lambda	0.53	4.52	4	20	0.009

Table 10. All teachers–Pre-/post-test comparisons

	Statistic	p
Task achievement	1,941.50	0.232
Coherence and cohesion	1,824.00	0.082
Lexical resource	1,950.00	0.241
Grammar and accuracy	1,384.00	< .001
Overall score	2,065.00	0.647

Table 11. Multivariate tests for all teachers

	Test	Value	F	df1	df2	p
Post (EG=1 & CG=2)	Pillai's trace	0.10	2.78	4	101	0.031
	Wilks' lambda	0.90	2.78	4	101	0.031
Grammar and accuracy	Pillai's trace	0.04	1.00	4	101	0.409
	Wilks' lambda	0.96	1.00	4	101	0.409

Teacher 4

Table 8 presents the pre-test scores of experimental and control groups on the dimensions of the test administered to students after the traditional feedback and Write & Improve feedback. Since there was no normality, referring to **Table 1**, Mann-Whitney U test was employed.

As seen in **Table 8**, control and experimental groups are not initially equal on TA (p = .017), and LR (p = .002). Only TA is used as covariate in the following MANCOVA analyses (**Table 9**) because TA, and LR had significant correlation between them (r = 0.76).

Table 9 indicates that when TA is used as a covariate, we found a significant overall difference in the means across multiple dependent variables simultaneously (p = .009). In other words, there is an effect of Write & Improve tool's feedback, when compared to traditional feedback after the assessment, on students' scores on the dimensions of the test in the classes of the fourth teacher. To locate where this significant difference appears, we did follow up univariate tests, however, we did not find any significant differences between the dimensions of the test administered to students.

All Teachers

Table 10 presents the pre-test scores of all teachers' experimental and control groups on the dimensions of the test administered to students after the traditional feedback and Write & Improve feedback. Since there was no normality Mann-Whitney U test was employed.

As seen in **Table 10**, control and experimental groups are not initially equal on GA (p < .001). Thus, GA scores are used as covariate in the following MANCOVA analyses (**Table 11**).

Table 11 indicates that without the covariate, we found a significant overall difference in the means across multiple dependent variables simultaneously (p = .031). However, with the covariate there is no effect of Write & Improve tool's feedback, when compared to traditional feedback after the assessment, on students' scores on the dimensions of the test in the classes of all teachers.

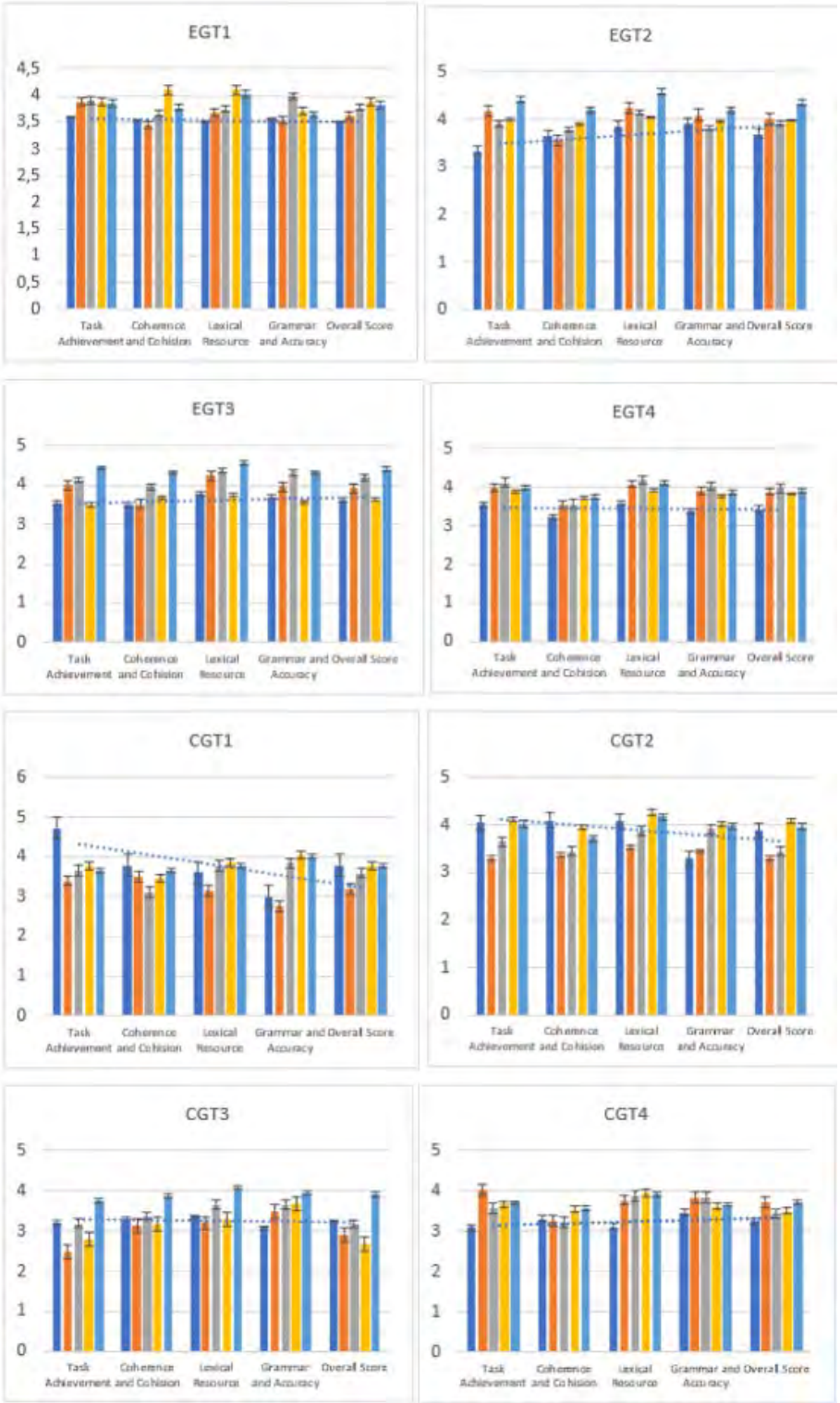

Figure 2. The change of students' scores over time (Source: Authors)

Longitudinal Effect

The effect of both types of feedback was assessed by comparing the post test scores of students in four different classes of teachers with MANCOVA. We also collected data from three assessments done during one semester. Along with the results of pre and post-tests, three assignments are presented in **Figure 2**. Each of the graphs (four experimental and four control groups) in **Figure 2** compromise five sets of data; TA, CC, LR, GA, and OS. Each of these sets contain bars for pre-test (blue), assignment 1 (orange), assignment 2 (grey), assignment 3 (yellow), and post-test (light blue) scores.

As seen in **Figure 2**, students' scores in four different teachers' classes, in two different feedback methods, and in five different dimensions of the test are waiving in all graphs. The trend line in each graph indicates the change over time. Either in the experimental group (first four graphs) or in control group (last four graphs) there is no consistent increase in the scores.

Summary of Findings

We tested the hypothesis that there is a difference between the effects of feedback given by teachers and the Write & Improve tool. The hypothesis was tested separately for four groups: control and experimental groups for each of the four teachers, as well as for the combined scores gathered from classes of all four teachers. We found no difference in the effects of both feedback methods, neither in individual teachers' classes nor in combined classes. The results of this study indicate that the Write & Improve tool is as effective as a teacher in providing feedback to students to improve their writing skills.

DISCUSSION

The investigation into the influence of AI-assisted writing tools on the writing skills of 1st year TFL students in Kazakhstan has yielded intriguing results. The students' writing skills were evaluated based on the IELTS writing assessment criteria, which include TA (the ability to fulfill the requirements of a task), CC (the connectedness of a text in paragraphs and as a whole text), and LR (the range of vocabulary used). Upon comparing pre and post-test scores for both control and experimental groups, it appears that there was no statistically significant difference between the impact of feedback from the Write & Improve tool and traditional teacher feedback across various dimensions of the test administered in the classes led by four different teachers (T1, T2, T3, and T4). Contrarily, in the case of T4, the outcomes, particularly concerning students' scores on various dimensions of the test, displayed a better result for Write & Improve compared to traditional teacher feedback. The intervention involved providing feedback on writing assignments using both Write & Improve and traditional methods. Write & Improve offered almost immediate feedback, consisting of suggestions for improvement, while traditional feedback required students to wait until instructors checked their work, which typically took about 5–7 days due to the workload of instructors. However, it's important to note that traditional instructor feedback was often more descriptive, as instructors could provide detailed explanations or examples if necessary, whereas Write & Improve did not offer such personalized feedback.

The data analysis indicates a lack of a significant difference between the effects of AI-assisted tool feedback and traditional teacher feedback. This finding stands in contrast to a number of prior studies exploring similar themes. In their discussion of traditional and contemporary methods, Hadžimehmedagić and Akbarov (2014) criticize the former, claiming that it is less student-centered and provide a list of disadvantages. This argumentation without any strong justification can be described as "weak arguments" that do not convincingly demonstrate the ineffectiveness of the traditional method compared to modern one. Furthermore, a significant weakness of this argument is that the effectiveness of the traditional method can vary depending on the nature and content of the class. As the saying goes, 'one size does not fit all,' suggesting that despite differences in delivery and instruction style, the traditional method can still be effective.

The findings of Schwerdt and Wuppermann (2011) on the role of teacher feedback in the development of learners' writing skills provide valuable insights into the positive impact of teacher feedback on enhancing learners' writing abilities. Specifically, Schwerdt and Wuppermann (2011) argue that constructive feedback from teachers contributes significantly to the improvement of learners' writing proficiency. This aligns with the argument being presented here, which emphasizes the importance of teacher feedback in fostering the development of writing skills among learners. However, it's important to note that while there is partial alignment between the findings of Schwerdt and Wuppermann (2011) study and the argument being made here, nuances and differences may exist in the specific contexts or methodologies employed in each study. Moreover, our findings align with the conclusions drawn by Taskiran and Goksel (2022) in their research on the effectiveness of AI-assisted feedback in the context of English as a Foreign Language (EFL) writing instruction. Taskiran and Goksel (2022) assert that AI-assisted feedback plays a beneficial role in enhancing EFL writing skills and can serve as a valuable supplementary tool for teachers. This supports the argument

being presented here, which emphasizes the potential of AI-assisted feedback as a valuable resource in language learning environments.

Moreover, by building upon the insights of Zaghlool and Khasawneh (2023), current study further bolsters the understanding that AI tools like Write & Improve or Grammarly play a constructive role in the learning process. Zaghlool and Khasawneh (2023) assert that these tools possess the ability to streamline learning and enhance its efficacy. This aligns with the argument presented in our research, underscoring the pivotal role of AI tools in facilitating and enriching the educational journey for students.

One of the primary focuses of the literature review is on the impact of Write & Improve on students' writing skills within the Kazakhstani context. Studies such as Karpova (2020) and Wali and Huijser (2018) provide positive reviews of the tool, citing its user-friendliness and effectiveness in improving writing skills and fostering student motivation. In the context of Kazakhstan, the findings of this review shed light on an important consideration: the usability of Write & Improve among Kazakhstani teachers. Contrary to the positive assessments provided by previous studies, this review highlights feedback from Kazakhstani teachers indicating a lack of user-friendliness attributed to the tool's limited functionality. This aspect underscores the significance of considering contextual factors, such as technological infrastructure and user preferences, when evaluating the effectiveness of AI-assisted feedback tools in diverse educational settings.

Although the research did not uncover any notable gap between the two feedback methods on the whole, it is prudent to explore potential variances in the caliber of feedback offered by each approach. Based on the findings of this study, certain teachers expressed reluctance to utilize AI-assisted tools citing concerns regarding ethical implications, while others emphasized the importance of delineating specific instructional purposes to maximize the effectiveness of such tools. Likewise, Hattie and Timperley (2007) have suggested that the specificity, timeliness, and personalization of feedback can markedly influence its efficacy. Thus, analyzing the content and delivery of feedback from both AI tools and teachers could contribute to their respective strengths and weaknesses. It is imperative to explore the nuances of this lack of significant difference and identify specific areas where AI tools and traditional teacher feedback demonstrated similarities or disparities. Gaining understanding into the factors that influenced the outcomes could help illuminate the interactions between technology and human-driven educational methods. Furthermore, it's important to note that this paper does not aim to conduct an in-depth analysis of the interviews conducted with teachers. Rather, the focus is on presenting the overarching themes and perspectives expressed by teachers regarding the utilization of AI-assisted tools in language teaching.

The findings of this study highlight the complementary role that AI tools like Write & Improve can play in educational settings, particularly in providing immediate, consistent, and accessible feedback to students. While AI tools may not replace the nuanced and context-specific feedback provided by experienced teachers, they offer distinct advantages in terms of availability and efficiency (Toscano, 2023). Studies have shown that students benefit from immediate feedback, as it promotes more effective learning by allowing learners to quickly correct their mistakes and reinforce correct usage (Karpova, 2021). Moreover, AI tools operate without the limitations of human fatigue or time constraints, making them especially useful in large classrooms or for students who need repeated practice (Curry & Riordan, 2021). Additionally, AI-driven feedback can alleviate some of the workload from teachers, enabling them to focus on more complex aspects of instruction such as critical thinking and content development (Taskiran & Goksel, 2022).

The findings of this study contribute to the ongoing discussion about the role of AI tools in educational settings, particularly in language learning and writing development. In line with theories of blended learning and technology-enhanced language learning, the use of AI tools like Write & Improve complements traditional instruction by offering immediate, consistent feedback that enables students to engage in self-regulated learning (Kukulska-Hulme & Shield, 2008). However, these tools are most effective when used as a supplement to teacher input rather than a substitute. The balance between AI and teacher feedback, as suggested by the results, supports frameworks that emphasize the integration of technology to enhance but not replace teacher expertise (Luckin, 2017).

Ethical considerations also play a critical role in the adoption of AI tools in education. Issues such as data privacy, bias in AI-generated feedback, and the depersonalization of learning experiences must be addressed to ensure responsible integration (Holmes et al., 2019). In the context of this study, while Write & Improve

offers several advantages–such as its availability and impartial feedback–it cannot provide the nuanced, context-specific insights that experienced teachers bring. Therefore, educators and policymakers should focus on creating guidelines for responsible AI use in classrooms, ensuring that it enhances rather than detracts from the personalized learning process (Chan & Hu, 2023).

For policymakers, the scalability of AI tools holds promise in addressing some of the resource constraints faced by educational systems, especially in large or under-resourced classrooms. The study's findings suggest that AI tools can alleviate teacher workload by handling more routine feedback tasks, allowing teachers to focus on higher-order instructional activities (Taskiran & Goksel, 2022). This has significant implications for educational equity, as AI tools can provide immediate, high-quality feedback to students who may not have regular access to one-on-one teacher support. However, any such adoption should be coupled with teacher training and ethical considerations to ensure effective, equitable implementation.

Conclusion, Limitation, and Implication

One limitation of this study is the relatively small number of participants in both the experimental and control groups. The small sample size may restrict the generalizability of the findings and limit the statistical power to detect significant effects (Gravetter et al., 2009, p. 656). While the findings provide valuable insights into the impact of AI-assisted writing tools on language learning outcomes within the specific context of our study, caution should be exercised when extrapolating these results to broader populations or different educational settings. It is important to acknowledge that the characteristics of our participants, such as their language proficiency levels, cultural backgrounds, and prior experiences with AI tools, may influence the outcomes observed in this study and may not be representative of all language learners. Therefore, while the findings contribute to our understanding of the potential benefits and limitations of AI-assisted writing tools, they may not be universally applicable. Future research with larger and more diverse samples is warranted to enhance the generalizability and reliability of the findings across various contexts and populations.

To sum up, while this study does not undermine the potential of AI writing tools such as Write & Improve, it underscores the effectiveness of instructors' feedback, which proves to be at least as impactful as AI-assisted feedback. Recognizing the competence of instructors and the rigorous selection process they undergo, coupled with the university's high rating from national accreditation agencies, further highlights the multifaceted factors influencing students' academic performance. Given the unique linguistic and cultural context of Kazakhstan, future research should explore how the findings of this study align with or diverge from existing language education practices in the country. Examining the compatibility of AI-assisted writing tools with pedagogical approaches offers valuable insights into their potential integration into language teaching methodologies. This examination enables educators to identify potential challenges, opportunities, and best practices for incorporating AI tools into classroom instruction in a manner that supports the achievement of learning objectives and enhances students' language learning experiences. Finally, it is crucial to consider the long-term effects of AI-assisted writing tools on students' writing development. While this study focused on short-term outcomes, longitudinal research could investigate whether the use of AI tools leads to sustained improvements in writing proficiency over time. Tracking students' progress and performance beyond the duration of the study could offer valuable insights into the lasting impact of AI-assisted feedback on language learning outcomes.

The hypothesis that there would be a difference between the effects of teacher feedback and Write & Improve tool feedback was not supported by the data. This finding held true across all individual teachers' classes as well as when considering the combined scores from all teachers' classes. Overall, the results suggest that, in the context of this study, feedback provided by the Write & Improve tool did not yield significantly different outcomes compared to traditional teacher feedback in terms of students' performance on the dimensions of the test assessed.

A key finding of this study is the lack of attention to ethical standards among students, particularly in the use of AI-driven tools like Write & Improve. Addressing this issue requires institutions to develop comprehensive educational programs and policies that promote ethical awareness and responsible use of technology in learning. One effective approach is the incorporation of ethics-focused courses or modules into the curriculum, specifically addressing issues such as academic honesty, plagiarism, and the ethical implications of AI tools in education. For instance, workshops or seminars that educate students about the

risks of over-reliance on AI tools and the importance of maintaining personal accountability in learning could be introduced (Ribble, 2011).

Institutions should also adopt clear policies on the ethical use of AI in academic work, detailing acceptable practices and the consequences of misuse. Creating an academic honor code that explicitly includes AI use and aligning it with broader ethical standards would help institutionalize responsible behaviors (Pavela, 1997). Furthermore, fostering a culture of integrity through student engagement–such as student-led ethics committees or peer-to-peer mentoring on academic integrity–could reinforce these values.

Policy changes could also include mandatory training sessions for both students and educators on digital literacy and AI ethics. Such training would not only raise awareness but also equip both groups with the tools to navigate the ethical challenges posed by emerging technologies in education (Holmes et al., 2019). By embedding ethical principles into both the academic framework and the institutional culture, schools can better prepare students to act responsibly in an increasingly technology-driven world.

Author contributions: AB: conducted the experiment and collected the data; **SA:** performed the literature review; **TZ:** contributed to the discussion and conclusions sections; **DG:** developed the methodology and edited the language; **NB:** supervised the research and handled the revisions. All authors approved the final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Ethics declaration: This study was approved by the SDU University Research Committee (approval registered in Protocol #6, dated 29.11.2023). Participants of the research have signed consent forms. All data has been stored on Bodaubekov Abbas' personal hard drive, so only the author has access to the information. During the data analysis phase, all personal information has been coded so that other authors do not have access to personal data.

Declaration of interest: The authors declare no competing interest.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

Alammari, A., & Abdel-Reheem Amin, E. (2023). EFL students' perception of using AI paraphrasing tools in English language research projects. *Arab World English Journals, 14*(3). https://doi.org/10.24093/awej/vol14no3.11

Almaiah, M. A., Alfaisal, R., Salloum, S. A., Hajjej, F., Thabit, S., El-Qirem, F. A., Lutfi, A., Alrawad, M., Al Mulhem, A., Alkhdour, T., Awad, A. B., & Al-Maroof, R. S. (2022). Examining the impact of artificial intelligence and social and computer anxiety in e-learning settings: Students' perceptions at the university level. *Electronics, 11*(22), Article 3662. https://doi.org/10.3390/electronics11223662

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*, Article 43. https://doi.org/10.1186/s41239-023-00411-8

Cheng, G. (2017). The impact of online automated feedback on students' reflective journal writing in an EFL course. *The Internet and Higher Education, 34*, 18–27. https://doi.org/10.1016/j.iheduc.2017.04.002

Chounta, I. A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers' perceptions of Artificial Intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education, 32*(3), 725–755. https://doi.org/10.1007/s40593-021-00243-5

Curry, N., & Riordan, E. (2021). Intelligent CALL systems for writing development: Investigating the use of Write & Improve for developing written language and writing skills. In K. Kelch, P. Byun, S. Safavi, & S. Cervantes (Eds.), *CALL theory applications for online TESOL education* (pp. 252–273). IGI Global. https://doi.org/10.4018/978-1-7998-6609-1.ch011

de los Ángeles Domínguez-González, M., Hervás-Gómez, C., Díaz-Noguera, M. D., & Reina-Parrado, M. (2023). Attention to diversity from artificial intelligence. *The European Educational Researcher, 6*(3), 101–115. https://doi.org/10.31757/euer.633

Fitria, T. N. (2021). Artificial intelligence (AI) in education: Using AI tools for teaching and learning process. In *Proceedings of the National Seminar* (pp. 134–147).

Gravetter, F. J., Forzano, L. A. B., & Rakow, T. (2009). Research methods for the behavioral sciences. Cengage Learning.

Hadžimehmedagić, M., & Akbarov, A. (2014). Traditional vs modern teaching methods. Advantages and disadvantages. *CORE*. https://core.ac.uk/download/pdf/153448544.pdf

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hervás Gómez, C., & Toledo Morales, P. (2018). Real and ideal perception of the intelligent classroom environment of future teachers. *Journal of Research in Science, Mathematics and Technology Education, 1*(1), 91–111. https://doi.org/10.31756/jrsmte.115

Heydari, M. & Marefat, F. (2020). Integrating computer- and teacher-provided feedback in an EFL academic writing context. In L. McCallum, & Coombe (Eds.), *The assessment of L2 written English across the MENA Region* (pp. 297–323). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-53254-3_13

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education promises and implications for teaching and learning.* Center for Curriculum Redesign.

Hussain, I. (2020). Attitude of university students and teachers towards instructional role of artificial intelligence. *International Journal of E-Learning and Distance Education, 5*, 158–177. https://doi.org/10.36261/ijdeel.v5i2.1057

Karpova, K. (2021). Integration of "Write and Improve" AWE tool into EFL at higher educational establishment: case study. *Celtic: A Journal of Culture English Language Teaching Literature & Linguistic, 7*, 137–150. https://doi.org/10.22219/celtic.v7i2.14036

Kompa, J. S. (2012). Disadvantages of teacher-centered learning. *Joana Stella Kompa*. https://joanakompa.com/2012/06/25/the-key-disadvantages-of-teacher-centered-learning/

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL, 20*(3), 271–289. https://doi.org/10.1017/S0958344008000335

Luckin, R. (2017). Towards artificial intelligence-based assessment systems. *Nature Human Behaviour, 1*(3), Article 0028. https://doi.org/10.1038/s41562-016-0028

Luckin, R., & Holmes, W. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

McGrath, C., Pargman, T. C., Juth, N., & Palmgren, P. J. (2023). University teachers' perceptions of responsibility and artificial intelligence in higher education-An experimental philosophical study. *Computers and Education: Artificial Intelligence, 4*, Article 100139. https://doi.org/10.1016/j.caeai.2023.100139

Nazaretsky, T., Cukurova, M., & Alexandron, G. (2022). An instrument for measuring teachers' trust in AI-based educational technology. In *Proceedings of the 12th International Learning Analytics and Knowledge Conference* (pp. 56–66). https://doi.org/10.1145/3506860.3506866

Niemi, H., & Niu, S. J. (2021). Digital storytelling enhancing Chinese primary school students' self-efficacy in mathematics learning. *Journal of Pacific Rim Psychology, 15*. https://doi.org/10.1177/1834490921991432

Pavela, G. (1997). Applying the power of association on campus: A model code of academic integrity. *Journal of College and University Law, 24*, 97–118.

Relmasira, S. C., Lai, Y. C., & So, C. F. H. (2021). Future jobs: Indonesian primary students' aspirations and teachers' predictions. *The European Educational Researcher, 4*(2), 209–225. https://doi.org/10.31757/euer.425

Ribble, M. (2015). *Digital citizenship in schools: Nine elements all students should know*. International Society for Technology in Education.

Robin B. R. (2008). Digital storytelling: A powerful technology tool for the 21st century classroom. *Theory Into Practice, 47*(3), 220–228. https://doi.org/10.1080/00405840802153916

Schwerdt, G., & Wuppermann, A. C. (2011). Is traditional teaching really all that bad? A within-student between-subject approach. *Economics of Education Review, 30*(2), 365–379. https://doi.org/10.1016/j.econedurev.2010.11.005

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Company.

Sit, C., Srinivasan, R., Amlani, A., Muthuswamy, K., Azam, A., Monzon, L., & Poon, D. S. (2020). Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: A multicentre survey. *Insights into Imaging, 11*, Article 14. https://doi.org/10.1186/s13244-019-0830-7

Sumakul, D. T. Y., Hamied, F. A., & Sukyadi, D. (2022). Artificial intelligence in EFL classrooms: Friend or foe? *LEARN Journal: Language Education and Acquisition Research Network, 15*(1), 232–256.

Taskiran, A., & Goksel, N. (2022). Automated feedback and teacher feedback: Writing achievement in learning English as a foreign language at a distance. *The Turkish Online Journal of Distance Education, 23*(2), 120–139. https://doi.org/10.17718/tojde.1096260

Timonen, P., & Ruokamo, H. (2021). Designing a preliminary model of coaching pedagogy for synchronous collaborative online learning. *Journal of Pacific Rim Psychology, 15*. https://doi.org/10.1177/1834490921991430

Toscano, M. A. (2023). *Write & Improve tool and writing skill* [Bachelor's thesis, Universidad Técnica de Ambato].

Väätäinen, J., & Ruokamo, H. (2021). Conceptualizing dimensions and a model for digital pedagogy. *Journal of Pacific Rim Psychology, 15*. https://doi.org/10.1177/1834490921995395

Wali, F., & Huijser, H. (2018). Write to improve: Exploring the impact of an automated feedback tool on Bahraini learners of English. *Learning & Teaching in Higher Education: Gulf Perspectives, 15*(1), 14–34. https://doi.org/10.18538/lthe.v15.n1.293

Wang, Y., Shang, H., & Briody, P. (2013). Exploring the impact of using automated writing evaluation in English as a foreign language university students' writing. *Computer Assisted Language Learning, 26*(3), 234–257. https://doi.org/10.1080/09588221.2012.655300

Yau, K. W., Chai, C. S., Chiu, T. K., Meng, H., King, I., & Yam, Y. (2023). A phenomenographic approach on teacher conceptions of teaching artificial intelligence (AI) in K-12 schools. *Education and Information Technologies, 28*, 1041–1064. https://doi.org/10.1007/s10639-022-11161-x

Yilmaz, H., Maxutov, S., Baitekov, A. & Balta, N. (2023). Student's perception of ChatGPT: A technology acceptance model study. *International Educational Review, 1*(1), 57– 83. https://doi.org/10.58693/ier.114

Zaghlool, Z. D., & Khasawneh, M. A. S. (2023). Incorporating the impacts and limitations of AI-driven feedback, evaluation, and real-time conversation tools in foreign language learning. *Migration Letters*, *20*(7), 1071–1083.