OPEN PRAXIS

The Effect of Feedback on Informative Text Writing: AI or Teacher?

RESEARCH ARTICLE

MEHMET ALI BAZ
SEVIL HASIRCI AKSOY
*Author affiliations can be found in the back matter of this article

INTERNATIONAL COUNCIL FOR OPEN AND DISTANCE EDUCATION

ABSTRACT

This study was conducted to examine the effects of feedback provided by a generative AI chatbot trained by the researcher and feedback given by a teacher on students' achievement and retention in informative text writing.In this context, a convergent parallel mixed-methods approach was adopted. The research was conducted with two groups at the fifth-grade middle school level, comprising experimental and control groups. For eight weeks, the chatbot provided feedback on the texts of students in the experimental group through the researcher, while the teacher provided feedback on the texts of students in the control group. To measure students' informative text writing performance, the Informative Text Writing Performance Assessment Scale developed by Hamzadayı and Dölek (2022) was utilized. Findings indicated that feedback provided by both the AI chatbot and the teacher contributed to improvements in students' informative text writing performance. However, it was determined that such improvement persisted significantly only within the teacher-feedback group. Based on this finding, it was concluded that while chatbot feedback was as effective as teacher feedback in improving writing performance, it did not exhibit comparable long-term retention effects. When comparing the content of feedback provided, it was observed that the chatbot's feedback was quantitatively more extensive than that of the teacher. Conversely, teacher feedback was superior regarding affective components and question-oriented feedback. Consequently, generative AI tools hold significant potential for reducing teachers' cognitive workload in writing instruction by providing supplementary and innovative feedback methods.

CORRESPONDING AUTHOR:
Sevil Hasirci Aksoy
Gaziantep University, Türkiye
sevilhasirci@gmail.com

KEYWORDS:
Artificial Intelligence; Writing Education; Feedback; Turkish Language Education; Middle School Students; Writing Performance; Informative Text; Chatbot; AI Feedback; Teacher Feedback

TO CITE THIS ARTICLE:
Baz, M. A., & Hasirci Aksoy, S. (2025). The Effect of Feedback on Informative Text Writing: AI or Teacher? *Open Praxis*, *17*(3), pp. 611–631. DOI: https://doi.org/10.55982/openpraxis.17.3.871

INTRODUCTION

Feedback is the reinforcing or corrective information provided to students regarding their performance in a learning environment (Bardwell, 1981). In writing instruction, formative feedback plays a crucial role in enhancing students' writing proficiency (Graham et al., 2011; MacArthur, 2016; Saliu-Abdulahi et al., 2017). Formative feedback not only helps students understand their current level of writing achievement and guides them on how to improve but also encourages continuous development throughout the learning process (Balta & Türel, 2013; Çetinkaya & Yolcusoy, 2020; Graham et al., 2012; Timperley & Parr, 2010).

Although teacher feedback is preferred by students (Çetinkaya & Hamzadayı, 2015) and often more effective than peer and self-feedback (Beach & Friedrich, 2006), the effort required to provide feedback significantly increases teachers' workload (Ballet & Kelchtermans, 2009; Dibbon, 2004). Additionally, given that teachers typically instruct multiple large classes, the time required for providing feedback may limit their ability to focus on teaching fundamental writing skills (Applebee & Langer, 2011; Graham, 2019). Reducing teachers' workload related to feedback provision in writing instruction could allow them to allocate more time to other aspects of the writing process (Biber et al., 2011; Graham et al., 2015; Grimes & Warschauer, 2010; Warschauer & Grimes, 2008; Wilson & Czik, 2016; Wilson & Roscoe, 2020). Consequently, alternative methods of providing feedback in writing education have become increasingly necessary.

Research and advancements in writing instruction have led to a broader range of feedback sources available to students. In addition to teacher feedback, peer feedback, self-assessment, writing workshops, oral conferences, and computer-assisted feedback have been incorporated into writing instruction (Hyland & Hyland, 2006). One such computer-assisted feedback method, automated writing evaluation (AWE), has attracted attention from educators due to its potential to assess student writing, assist teachers, and reduce their feedback workload (Kellogg et al., 2010; Liu et al., 2024; Wilson & Roscoe, 2020; Wilson et al., 2022). These systems generally utilize natural language processing (NLP) techniques or artificial intelligence (AI) to evaluate students' written texts. Numerous studies have demonstrated that feedback provided through such systems positively impacts students' writing performance (Graham et al., 2015; Grimes & Warschauer, 2010; Lee, 2019; Roscoe et al., 2017; Stevenson & Phakiti, 2014; Wilson & Czik, 2016; Wilson & MacArthur, 2016; Wilson & Roscoe, 2020; Zhai & Ma, 2022).

On the other hand, some researchers (MacArthur, 2016; Wilson et al., 2022; Yancey et al., 2004) have argued that AWE applications predominantly focus on scoring rather than providing meaningful feedback, and due to the variability in students' responses to such feedback, these tools may be detrimental to students' writing skills. Additionally, they have emphasized that integrating these tools into learning environments takes time (Chen et al., 2022; Moore & MacArthur, 2016; Shermis, 2014; Ware & Warschauer, 2006; Xie et al., 2020). While these discussions continue, the use of both paid and free AWE applications has been increasing, particularly in technologically advanced countries like the United States and China. However, while studies suggest that these applications offer scoring comparable to—or even higher in quality than—that of teachers (Shermis & Hamner, 2013; Stevenson & Phakiti, 2014), their long-term effectiveness in improving students' writing skills remains highly limited. The available evidence is insufficient, and researchers have emphasized the need for more comprehensive studies on this topic (Xiao et al., 2025).

Amid ongoing debates about the advantages and drawbacks of integrating AI-based systems into education, OpenAI's 2022 release of ChatGPT has sparked both excitement (Derakhshan & Ghiasvand, 2024; Lee, 2023) and concern (Akgun & Greenhow, 2022; Chomsky et al., 2023; Harari, 2023). The fundamental distinction between ChatGPT and similar models (Claude, Poe Ai, Gemini, Grok, Bing, etc.) and traditional AI-driven autonomous systems is their ability to not only mimic human language but also process and generate text, comprehend generated content, ask and answer questions, and interact dynamically with users (Bozkurt, 2020, 2023a). The introduction of generative AI (GenAI) tools has not only increased the popularity of automated written feedback (AWF) but also heightened expectations regarding the quality of feedback these systems can offer (Barrot, 2023; Burstein, 2020). Based on these developments, AWF shows significant potential for providing feedback in writing instruction, and with advancements in GenAI models, educators are likely to incorporate these systems

more frequently into their teaching. Additionally, some studies (Hsieh et al., 2017; Kluger & Adler, 1993) have classified AWF as a form of computer-assisted feedback, similar to AWE, that includes both automated feedback and GenAI responses.

With the introduction of GenAI tools, AI-supported chatbots are increasingly being used to provide feedback on students' writing (Guo & Wang, 2024; Kohnke et. al., 2023; Xiu-Yi, 2024; Yang & Kyun, 2022; Yang & Li, 2024). Additionally, GenAI platforms such as Poe AI enable users to create AI-powered chatbots trained for specific purposes, using a foundational GenAI model, at no cost.

Despite the growing body of research on the use of AWE and AWF in writing education—particularly in English and Chinese—an extensive literature review reveals no applied research on this subject in Turkish writing instruction. A previous review suggested that AI could be utilized to provide feedback to students (Akkaya & Çıvgın, 2021); however, no empirical studies have been conducted in the Turkish context. Therefore, the present study aims to examine the impact of feedback provided by a GenAI chatbot developed by the researchers on the informative text writing performance and retention of middle school students. Furthermore, the study seeks to analyze the nature, similarities, and differences of the feedback provided. This study is significant, as it represents the first applied research on the use of GenAI systems for feedback provision in Turkish writing instruction. Moving forward, this research is expected to support the development of GenAI-based AWE and AWF systems, particularly in the context of native language education, and further advance feedback-driven AI applications in writing pedagogy.

RESEARCH AIM

The aim of this study is to examine the impact of feedback provided by a teacher and a chatbot on informative texts written by fifth-grade students in Turkish language courses. Specifically, the study investigates how the content of the feedback influences students' writing performance in informative text composition.

In line with this objective, the research seeks to answer the following questions:

1. Is there a significant difference in the effect of formative feedback provided by a chatbot versus a teacher on students' performance in writing informative text?
2. Does the effect of formative feedback from a chatbot or a teacher on students' informative text writing performance persist over time?
3. What are the similarities or differences between the formative feedback provided by the chatbot versus the teacher?

LITERATURE

TYPES OF FEEDBACK

Although the literature presents various classifications of teacher feedback on written texts, it is generally categorized into feedback type and feedback level. Feedback types are further classified into direct and indirect feedback based on how they are presented to the student (Biber et al., 2011). Direct feedback provides explicit instructions for revising the text, while indirect feedback requires students to identify and correct their own mistakes based on cues provided in the feedback. Feedback can include questions, informative elements, or praise (Cho et al., 2006; Nelson & Schunn, 2009; Peterson et al., 2004; Shute, 2008). Feedback level refers to the teacher's focus when providing feedback, which can be categorized into five areas: vocabulary, grammar, mechanics, organization, and content (Biber et al., 2011). In their study on the effects of automated assessment and teacher feedback on student motivation and writing quality, Wilson and Czik (2016) classified feedback types as directive, questioning, informative, or praising, and feedback levels into lower-order writing skills (spelling, capitalization, punctuation, sentence structure, grammar, formatting, and word choice) and higher-order writing skills (ideas and elaboration, organization, style, and self-reflection).

Studies at the K–12 level have shown that teacher feedback plays a significant role in improving students' writing skills. However, its effectiveness varies depending on the type and level of feedback. Direct feedback is generally preferred by both students and teachers

(Aridah et al., 2017) as it enhances clarity and comprehensibility (Wahyudi et al., 2024). Notably, teachers often provide more feedback than students expect (Aridah et al., 2017), which can sometimes lead to confusion and anxiety (Wahyudi et al., 2024). Teachers tend to focus on standardizing the written text rather than improving its content or organization (Matsumura et al., 2008). Regardless of the type and level, feedback involves corrections, and students may respond by asking questions, accepting, or ignoring the feedback (Rismawati & Hikmat, 2018). To maximize the effectiveness of feedback in writing instruction, teachers should carefully consider both the content and delivery of their feedback (Wahyudi et al., 2024).

In conclusion, the method and delivery of feedback play a critical role in enhancing students' writing performance and quality. Therefore, feedback should be systematically analyzed in terms of its source and effectiveness.

STUDIES ON FEEDBACK

Feedback in writing instruction is typically analyzed from four main perspectives: teacher feedback, peer feedback, self-feedback, or computer-assisted feedback (Dawson et al., 2023; Loncar et al., 2023). The following section outlines key studies associated with each approach.

Teacher feedback plays a crucial role in enhancing students' writing skills, offering revision opportunities that help them grow into proficient writers (Jerry et al., 2013). Both teacher and peer feedback contribute positively to students' writing performance and their attitudes toward the writing process (Ghani & Asgher, 2012). Teacher feedback can be either direct or indirect, with its effectiveness depending on how well it addresses student needs and classroom realities (Srichanyachon, 2012). However, teachers sometimes ignore students' feedback preferences and perspectives, which can hinder the learning process (Raihany, 2014). To maximize the benefits of feedback, teachers should provide clear explanations and incorporate positive comments to enhance student motivation (Srichanyachon, 2012). Teacher feedback can also be complemented by feedback from peers, which offers immediate clarification and alternative perspectives, fostering genuine communication among students (Ghani & Asgher, 2012).

Recent research on teacher feedback in writing instruction highlights its importance in the cognitive, affective, and psychomotor dimensions of writing. Mamad and Vigh's (2024) study examined teacher and student perceptions and practices regarding written feedback, revealing a need for further research on the relationships and differences between these perspectives. One identified gap involves the disconnect between fieldwork and theoretical knowledge in preservice writing teacher training (Peck & Kavanagh, 2024). In second-language writing, feedback from multiple sources has shown the greatest impact on student motivation (Cen & Zheng, 2024). For university entrants, both direct and indirect corrective feedback enhances writing performance, with direct feedback combined with metalinguistic explanations proving particularly beneficial (Wondim et al., 2024). These studies highlight the complex nature of teacher feedback in writing instruction and its potential to enhance student achievement at various levels.

In K–12 writing education, peer feedback has shown more positive outcomes than teacher feedback in some cases, and can be as effective as teacher feedback in improving students' writing performance (Cui et al., 2021; Gielen et al., 2010). When combined with teacher feedback, peer feedback can significantly enhance the quality of writing across multiple drafts (Pojslová, 2022). However, students often require structured guidance to provide high-quality feedback. Teaching guides and examples can improve the comprehensiveness of peer feedback, but students may still struggle to offer actionable suggestions (Gyamfi et al., 2024). This raises questions regarding the multisource functionality of peer feedback. A meta-analysis study found that multisource feedback had the most substantial motivational effect on L2 writing, followed by single-source feedback, including peer feedback (Cen & Zheng, 2024). In practice, peer review has been shown to significantly improve students' descriptive writing scores compared to those who did not receive peer feedback (Angraini, 2023). While these studies highlight the potential of peer feedback in writing instruction, they also emphasize the need for structured guidance to maximize its effectiveness.

Self-feedback in writing instruction has been examined alongside peer and teacher feedback, revealing significant effects on student outcomes. While self-feedback can contribute to writing development, peer and teacher feedback have proven more effective in fostering

collaborative writing skills. However, when combined with peer feedback, self-feedback has shown notable benefits in promoting critical thinking and increasing both behavioral and emotional engagement in learning (Zou et al., 2022). Research suggests that the application of feedback, rather than the quantity of feedback, plays a more significant role in supporting writing development (Lu et al., 2021).

Research on computer-mediated feedback in writing instruction has demonstrated its effectiveness in both feedback quality and teacher workload reduction. AWE systems can provide effective feedback while alleviating some of the burden on teachers (Fan & Ma, 2022). Wilson and Roscoe (2020) found that AWE feedback increased students' writing self-efficacy and performance on language skill assessments, while Sarı & Han (2022) found that combined automated tutor feedback was as effective as full tutor feedback in improving analytical writing scores and more effective in reducing grammatical and mechanical errors. However, feedback's effectiveness may depend on factors such as feedback type and students' analytical language abilities. For example, in a group receiving direct metalinguistic feedback, language analytic ability exhibited a stronger correlation with acquisition compared to a group receiving only direct feedback (Sheen, 2007). These studies suggest that when implemented appropriately, AWE systems can serve as valuable tools in writing instruction.

Recent studies have obtained varying results regarding the effectiveness of AWF. A systematic review of 83 studies examining AWF's performance, students' perceptions, usage, engagement, and the factors influencing these aspects found that AWF has both positive, negative, and neutral effects on advanced-level language and writing courses (Shi & Aryadoust, 2024). A meta-analysis of 20 studies found that AWF has a moderate overall effect on students' writing performance ($g = 0.55$) but also highlighted significant heterogeneity in the results (Fleckenstein et al., 2023). In the context of open and distance English language instruction, a 9-week quasi-experimental study investigated the impact of AWF. During the first 3 weeks, students received feedback from an English instructor on their writing tasks, and in the last 3 weeks, they received feedback from an automated system. Pre- and posttest results indicated that both teacher and automated feedback contributed to improvements in academic writing proficiency, with students in the teacher feedback group achieving slightly higher writing performance scores (Taskıran & Göksel, 2022). These findings suggest that AWF can serve as a valuable tool in writing instruction, potentially reducing teacher workload while providing students with timely and personalized feedback.

The literature review highlights the significant role of feedback in writing education and its positive impact on student achievement. As technological advancements and evolving educational paradigms shape feedback methods, alternative approaches like GenAI–based applications continue to emerge. Given these developments, the integration of diverse feedback mechanisms remains crucial for optimizing writing instruction.

METHODOLOGY

This study aims to examine the impact of feedback provided by a teacher and a chatbot on the informative texts written by fifth-grade students in Turkish language courses. Specifically, the study investigates how feedback influences students' writing performance and analyzes the content of the feedback.

To achieve this, the study adopts a convergent parallel mixed-methods design, which involves collecting, analyzing, and integrating both quantitative and qualitative data to provide a comprehensive understanding of the research problem. This approach strengthens the study by leveraging the advantages of both data types to obtain reliable and valid results. In this mixed-methods design, quantitative and qualitative data are collected simultaneously but analyzed separately, and the findings are later compared and synthesized (Creswell, 2020).

For the quantitative component, the study employs a quasi-experimental design. Compared to traditional experimental methods, quasi-experiments are better suited for educational environments, where naturally existing classroom groups are used instead of artificially created groups (Baştürk, 2014; Eckhardt & Ermann, 1977). Since forming artificial experimental groups is not feasible in school settings, a pretest, posttest, and follow-up test design was implemented using two intact groups: One group was designated as the experimental group, which received

chatbot feedback. The other was the control group, which received teacher feedback. Both groups underwent pretests to ensure comparability before the intervention. The experimental and control groups were then randomly assigned (Creswell, 2020; Patton, 2002).

QUALITATIVE DATA COLLECTION

To analyze the qualitative aspect, a content analysis method was applied to the feedback given by both the teacher and the chatbot on students' texts. Content analysis systematically processes written data by categorizing, coding, summarizing, and reporting it (Cohen et al., 2007; Flick, 1998; Mayring, 2004). This approach enables a deeper understanding of the nature, similarities, and differences in the feedback provided by the teacher and the chatbot.

STUDY GROUP

In experimental studies, random assignment of participants to groups is ideal. However, when random assignment is not feasible, researchers may use preexisting groups that are relatively similar in key characteristics (Büyüköztürk et al., 2023). For this study, the sample consists of fifth-grade students from a public middle school in Gaziantep, Turkey. Both groups were selected from the same school to minimize external variables. After obtaining the necessary ethical approval, a pretest was administered to both groups to determine whether there were any significant initial differences in writing skills. The pretest revealed no significant differences between the groups [t: −0.824, p: 0.414, $p > 0.05$]. Following the pretest, participants were randomly assigned to either the experimental (chatbot feedback) or control (teacher feedback) group. This quasi-experimental approach ensures that comparisons between the groups remain valid and reliable, maintaining the integrity of the research findings.

DATA COLLECTION PROCESS

This study first examined research on GenAI and feedback. A literature review indicated that no chatbot had been used to provide feedback on informative texts written by fifth-grade students in Turkish language courses. Therefore, a feedback chatbot was developed using the Claude-3-Haiku foundational model within the Poe AI platform, which enables users to connect with various GenAI tools and create customized chatbots for free. (For the chatbot's interface and access link, see Appendix 1). This chatbot was specifically trained to provide formative feedback on informative texts written by fifth-grade students in the Turkish language course. For this purpose, a clear and explicit command (prompt) was designed to guide the chatbot's responses. Additionally, eight academic sources related to formative and high-quality feedback were uploaded to the system to enhance its knowledge base (Coşkun & Tamer, 2015; Graham et al., 2015; Hamzadayı & Dölek, 2022; Ministry of National Education [MEB], 2019; Sarıkaya, 2023; Ülper, 2019; Wisniewski et al., 2020). Based on the subsequent tests and expert opinions, the final version of the chatbot command was determined (see Appendix 2).

Following the refinement of the chatbot, an 8-week implementation was conducted in a fifth-grade class of a public middle school under the MEB. Before the implementation, both the experimental and control groups were assigned a writing task based on a theme from the MEB Turkish language textbook. Students were given two class periods to write an informative text on the selected topic. The texts were evaluated using the Informative Text Writing Achievement Scale developed by Hamzadayı and Dölek (2022). For 8 weeks, both groups were assigned the same topics, selected based on the theme and content of the textbook. Students were required to write an informative text on the given topic within 2 class hours. In the control group, the teacher provided written feedback on these texts, whereas in the experimental group, written feedback was provided by the chatbot developed by the researcher.

At the end of the 8-week period, students in both groups were once again asked to write an informative text within 2 class hours on the same topic assigned before the intervention. These texts were reassessed using the same evaluation scale applied at the beginning of the study. To assess retention, 2 months after the completion of the experimental process, students were again asked to write an informative text on the same topic within 2 class hours. These texts were also evaluated using the same scale. The list of topics assigned to students during the writing instruction process is provided in Appendix 3.

DATA COLLECTION TOOLS

1. INFORMATIVE TEXT WRITING ACHIEVEMENT ASSESSMENT SCALE (FOR QUANTITATIVE DATA COLLECTION)

This scale was developed by Hamzadayı and Dölek (2022) to assess the informative text writing performance of middle school, high school, and university students. To determine the content validity index (CVI) of the scale items, expert opinions from 10 specialists were obtained. To establish the reliability of the scale using Kendall's coefficient of concordance (W), informative texts from a total of 63 students (21 middle school, 21 high school, and 21 university) were collected and evaluated by three raters. The analysis revealed that the CVI values of the scale items and the inter-rater Kendall's W coefficient were greater than 0.80. These findings indicate that the Informative Text Writing Achievement Rubric, consisting of five sub-dimensions—format, content development, content organization, word choice, and grammar—and 22 items, is both valid and reliable. For this study, the first two optional items under the format dimension were excluded, as they did not align with the research objective (see Appendix 4).

2. WRITTEN FEEDBACK PROVIDED BY THE TEACHER AND CHATBOT (FOR QUALITATIVE DATA COLLECTION)

A total of 100 written feedback instances (50 from the teacher and 50 from the chatbot) provided to fifth-grade students over 8 weeks were used as qualitative data sources for this research.

DATA ANALYSIS

NORMALITY TEST

The normality distribution of the data was first examined, using the Shapiro–Wilk test due to the sample size. The Shapiro–Wilk test is one of the most powerful normality tests, particularly for samples of less than 50 (Büyüköztürk, 2007; Razali & Wah, 2011).

As the p-values for the pretest and posttest scores of both the experimental and control groups were greater than 0.05, it was determined that the scores did not show a significant (extreme) deviation from a normal distribution. This indicates that statistical tests requiring the normality assumption can be applied. Additionally, skewness values within the range of ±1 and kurtosis values between −2 and +2 (George & Mallery, 2010) further confirmed that the pretest and posttest scores followed a normal distribution. When the data conform to a normal distribution, parametric tests—although requiring certain assumptions about distribution—are more robust in accurately representing central tendency and distribution properties (Field, 2013; Tabachnick & Fidell, 2013). However, when examining the follow-up test, the Shapiro–Wilk p-value for both groups was found to be less than 0.05. Despite the skewness and kurtosis values, the p-value of the follow-up test required the use of a nonparametric test.

TESTS USED IN DATA ANALYSIS

Depending on the normal distribution of the data:

- A paired samples t-test was used for within-group pretest and posttest differences.
- An independent samples t-test was used to compare pretest and posttest differences between the two groups.
- A Friedman test was used to compare repeated measures of pretest, posttest, and follow-up test scores.

COHEN'S D AND R EFFECT SIZE

Since the data followed a normal distribution, Cohen's d was used to determine the effect size of AI-supported feedback on students' informative writing performance scores. Effect size provides a practical measure that complements statistical significance. It also helps assess

how meaningful the effect is in real-world applications. Unlike p-values, which indicate whether an effect is statistically significant, Cohen's d expresses the magnitude of the effect using a standardized measure (Cohen, 1988; Ellis, 2010; Sullivan & Feinn, 2012). The Wilcoxon signed-rank test effect size was calculated by dividing the Z statistic by the square root of the total sample size (Field, 2013), and Cohen's classification (Cohen, 1988) was used to interpret the effect size.

KENDALL'S W

Kendall's W coefficient of concordance is a nonparametric statistical method that measures the agreement level among multiple raters by ranking their evaluations of a group. This coefficient specifically works with ranking data and ranges between 0 and 1: W = 0 indicates no agreement, while W = 1 represents perfect agreement. This method evaluates how consistently different raters rank the same sample set (Kendall, 1938). Kendall's W is a nonparametric test, meaning that it does not require a normal distribution, making it a flexible statistical tool (Field, 2013). In this study, Kendall's W coefficient was calculated using scores obtained from four main criteria and 20 sub-criteria for evaluating informative texts. The coefficient was found to be 0.83, indicating that the agreement among raters was sufficiently high (Howell, 2002).

QUALITATIVE DATA ANALYSIS

AI and teacher feedback texts were analyzed using systematic sampling to ensure representativeness and comparability. Through maximum variation sampling (Teddlie & Yu, 2007), every fourth text was selected, resulting in a total of 100 texts (50 AI, 50 teacher).

Both researchers divided each text into idea units according to feedback types identified in Cho et al. (2006), Wilson and Czik (2016), and Guo and Wang (2024). To ensure coding agreement, a coding scheme was developed (see Appendix 5), and 12 texts were analyzed together by the researchers. After discussing and resolving coding disagreements, the remaining 100 texts were categorized under two main themes: feedback types and written expression structure. Sample data related to the feedback types in the coding scheme were presented in the AI-generated and teacher-generated feedback columns, as well as in the researcher notes (see Appendix 5 and Appendix 8). Students were coded as Participant 1, Participant 2, and so on. The coding scheme included five types of feedback: informative, directive, questioning, praise, and summarizing. Additionally, each feedback type was classified under one of three main categories: content, organization, or language expression. The obtained data were analyzed using frequency and percentage distributions. In qualitative data analysis, ensuring the reliability of findings and maintaining inter-coder consistency is essential. According to the Miles and Huberman model, inter-coder agreement should be at least 80% (Miles & Huberman, 1994; Patton, 2002). In this study, inter-coder agreement was calculated at 83%, confirming that the qualitative data analysis was sufficiently reliable.

FINDINGS

This section presents the study's findings in line with the research questions. The first two subsections report statistical results on the effects and retention of AI- and teacher-provided formative feedback on students' informative writing. The third subsection analyzes similarities and differences in feedback content.

1. COMPARING EFFECTS OF CHATBOT AND TEACHER FEEDBACK ON INFORMATIVE TEXT WRITING PERFORMANCE

The comparison of the effect of AI-supported feedback between groups and within each group is presented in Tables 1 and 2.

	GROUP	N	MEAN	S	df	t	p	d
Post- test	Artificial intelligence	26	79,1923	14,13370	50	–1,428	0,160	0,396
	Teacher	26	74,3846	9,75326				

Table 1 Comparison between groups.

Table 1 shows that there was no statistically significant difference between the posttest scores of the AI and teacher feedback groups [t = −1.428, p > 0.05]. Although the mean posttest score of the AI group was slightly higher than that of the teacher feedback group, the magnitude of the difference between the two groups was also small (d = 0.396). This suggests that while there is no statistical significance, there may still be a meaningful effect.

PRE-TEST-POST-TEST		N		S	df	t	p	d
Artificial intelligence	Pre-test	26	45,4231	13,24590	25	−13,712	0,000	2,50
	Post-test	26	79,1923	14,13370				
Teacher	Pre-test	26	42,5000	12,32315	25	−12,185	0,000	2,49
	Post- test	26	74,3846	9,75326				

Table 2 In-group comparison.

A statistically significant difference was observed between the pretest and posttest scores of the group that received AI-supported feedback [t = −13.712, p < 0.05]. Additionally, since the Cohen's d value exceeded 1.0, it can be stated that AI-supported feedback had a very large effect (d = 2.50) on informative text writing performance. A statistically significant difference was also found between the pretest and posttest scores of the group that received teacher feedback [t = −12.185, p < 0.05], with the same level of effect size (d = 2.49). When analyzing the difference between the pretest and posttest in sub-variables, it was found that the mean of each variable increased in the posttest, and a significant difference (p = 0.000) with a large effect was observed (see Appendix 6 and Table 7).

2. COMPARING THE LASTING EFFECTS OF CHATBOT AND TEACHER FEEDBACK ON INFORMATIVE TEXT WRITING

The difference in retention between the AI feedback and teacher feedback groups is another subproblem.

		N	MEAN RANK	CHI-SQUARE	df	p	W
Artificial intelligence	Pre-test	26	1,90	3,571	2	0,168	0,0686
	Post- test	26	2,29				
	Follow-up test	26	1,81				
Teacher	Pre-test	26	1,40	39,864	2	0,000	0,7666
	Post- test	26	1,60				
	Follow-up test	26	3,00				

Table 3 Change over time.

When examining the rank averages of the AI feedback group, an increase was observed from the pretest to the posttest, but a decrease occurred in the follow-up test (Table 3). Additionally, based on the p-value, it was determined that there was no statistically significant difference between the measurements [p = 0.168, p > 0.05]. By contrast, in the teacher feedback group, the success average increased consistently from the pretest to the follow-up test, and a statistically significant difference was found [p = 0.000, p < 0.05]. Beyond the difference itself, the effect also varied: AI-supported feedback showed a small effect on informative text writing performance over time (W = 0.0686), whereas teacher-supported feedback had a large effect (W = 0.7666).

		N	MEAN RANK	SUM OF RANKS	Z	p	r
Artificial intelligence	Negative Ranks	18	11,03	198,50	−,970	,332	−0.199
	Positive Ranks	7	18,07	126,50			
	Ties	1	–	–			
Teacher	Negative Ranks	0	,00	,00	−4,459	0,000	−0.875
	Positive Ranks	26	13,50	351,00			
	Ties	0	–				

Table 4 Post-test-follow-up test difference.

The Wilcoxon signed-rank test conducted between the posttest and the follow-up test showed that while all participants in the teacher-supported instruction group showed a positive increase from the posttest to the follow-up test, in the AI-supported instruction group, a positive increase in informative text writing performance was observed in seven participants, whereas a decrease was observed in 18 participants, and one participant remained the same (Table 4). Therefore, although a significant difference was found between the posttest and the follow-up test in the teacher-supported group (p = 0.000), no significant difference was observed in the AI-supported experimental group (p = 0.332). Similarly, while a small effect size was found in the experimental group (r = 0.199), the effect size in the teacher-supported group was determined to be large (r = −0.875).

Table 5 Change over time in terms of sub-dimensions.

			N	MEAN RANK	CHI-SQUARE	df	p	W
Content creation	Artificial intelligence	Pre-test	26	1,00	43,429	2	0,000	0,835
		Post- test	26	2,31				
		Follow-up test	26	2,69				
	Teacher	Pre-test	26	1,10	34,940	2	0,000	0,672
		Post- test	26	2,63				
		Follow-up test	26	2,27				
Content organisation	Artificial intelligence	Pre-test	26	1,00	42,716	2	0,000	0,821
		Post- test	26	2,48				
		Follow-up test	26	2,52				
	Teacher	Pre-test	26	1,04	40,735	2	0,000	0,783
		Post- test	26	2,27				
		Follow-up test	26	2,69				
Word usage	Artificial intelligence	Pre-test	26	1,0	45,277	2	0,000	0,871
		Post- test	26	2,31				
		Follow-up test	26	2,69				
	Teacher	Pre-test	26	1,15	33,250	2	0,000	0,639
		Post- test	26	2,19				
		Follow-up test	26	2,65				
Grammar	Artificial intelligence	Pre-test	26	1,00	43,583	2	0,000	0,838
		Post- test	26	2,35				
		Follow-up test	26	2,65				
	Teacher	Pre-test	26	1,08	40,696	2	0,000	0,783
		Post- test	26	2,23				
		Follow-up test	26	2,69				

A comparison of sub-variables revealed that in the group receiving AI feedback, each variable increased from the pretest to the follow-up test, while in the teacher-supported group, there was an increase in the mean rank for every variable except content creation (Table 5). In both groups, a significant difference was observed between the pretest, posttest, and follow-up test for content creation, content organization, word usage, and grammar (p = 0.000), with a large effect size (W = 0.639 to 0.871). When comparing the posttest and follow-up test (see Appendix 6 and Table 8), no significant difference was found in the AI-supported group for content creation (p = 0.119) or content organization (p = 0.283), while a significant difference was observed in the teacher-supported group (content creation p = 0.002; content organization p = 0.016), with a large effect size. While a significant difference was found in word usage in both groups, the effect size was also deemed equal. No significant difference in grammar was found in either group, and the effect size was small. In sum, differences between the posttest and follow-up test were observed between the two groups; however, these differences did not affect the values between the pretest, posttest, and follow-up test.

3. COMPARING SIMILARITIES AND DIFFERENCES IN CHATBOT AND TEACHER FEEDBACK CONTENT

Teachers and AI provide different types of feedback and vary in the components of written expression, as seen in Table 6.

Table 6 Differences in feedback rate.

TYPE OF FEEDBACK	ARTIFICIAL INTELLIGENCE		TEACHER		COMPONENTS OF WRITTEN EXPRESSION	ARTIFICIAL INTELLIGENCE		TEACHER	
	f	%	f	%		f	%	f	%
Router	81	26,82	34	14,66	Language/Spelling and Grammar	54	21,52	62	33,16
Informative	140	46,36	126	54,31	Content	142	56,57	64	34,22
Question	0	0	7	3.01	Organisation	55	21,91	61	32,62
Praise	63	20,86	60	25,86	Total	251	100	187	100
Summariser	18	5,96	5	2,16					
Total	302	100	232	100					
Pearson Chi-Square	25,631				Pearson Chi-Square	3,225			
p	0,000				p	0,199			
Cramer's V	0,219				Cramer's V	0,082			

In terms of feedback type, informative feedback was the most frequently provided type in both groups. However, in the AI group, directive feedback ranked second, followed by praise and summarizing feedback. In contrast, praise was the second most common type in the teacher group. This indicates that while AI prioritizes providing information along with suggestions for changes, teacher feedback tends to motivate participants through praise after providing information. In terms of written expression components, AI focused more on content, whereas the teacher balanced all three components. AI tended to provide content-related feedback by combining information, guidance, and partial summarization. In contrast, teacher feedback was more likely to provide information on each component, offer praise related to these components, and occasionally prompt reflection through questions.

Analysis revealed a significant difference between the groups in feedback type ($p = 0.000$), with a moderate effect size. However, there was no significant difference between the two groups in written expression components, meaning the effect was very low. Although the teacher primarily provided informative feedback on content creation and content organization, AI's content-related feedback was more closely associated with praise and summarization. While this led to a significant difference in the posttest, it did not have a significant impact on the follow-up test.

DISCUSSION

According to the quantitative findings of this study, formative feedback provided by both the teacher and the chatbot had a similar effect on the informative text writing performance of fifth-grade students. A significant difference was observed between the pretest and posttest scores of both groups, suggesting that feedback from the chatbot was as effective as that provided by the teacher. This finding is supported by several studies (Alsofyani & Barzanji, 2025; Khan et al., 2025; Kim et al., 2024). However, a study by Steiss et al. (2024) indicated that teacher feedback was more effective. While the posttest results showed differences, the fact that teacher feedback resulted in a significant difference in the retention test, whereas AI feedback did not, aligns with the findings of Steiss et al. (2024). In their study, Steiss et al. (2024) examined the quality of formative feedback provided by both human evaluators and ChatGPT-3.5 on the argument-based writing of 200 randomly selected students from grades 6 to 12 across 26 different regions in Southern California. Feedback quality was assessed according to criterion-based evaluation, clarity of revision instructions, accuracy, prioritization of essential features, and the use of a supportive tone. The results indicated that human evaluators provided higher-quality feedback than ChatGPT in all categories except criterion-

based feedback. This suggests that the higher quality of teacher feedback had a greater impact on retention. Thus, the short-term performance gains observed when GenAI tools are used in educational settings—which could be described as a flash-in-the-pan effect—should not lead educators to become overly reliant on such tools. Instead, it is crucial to critically evaluate, observe, and carefully integrate these technologies into the learning process with an awareness of their long-term implications. Moreover, given that GenAI technologies and their applications are relatively new, future research with clearly defined standards and improved AI literacy may yield different results.

GenAI tools clearly hold potential for providing feedback on students' texts while teaching writing (Tlili et al., 2023). The findings of this study show that one such potential is that GenAI chatbots can serve as an effective and innovative tool for providing feedback in Turkish writing education. However, challenges such as limited access to GenAI tools, teachers' lack of necessary knowledge and skills to integrate these tools into the educational process, and negative attitudes toward AI adoption (Abdelhalim, 2023) hinder the realization of their full potential. This risk may lead to a widening learning gap between students who utilize AI tools and those who do not. Additionally, the pricing of GenAI tools, which varies based on their capabilities, could contribute to educational inequality between students with access to these tools and those without (McGee & Sadler, 2024). Even with equal conditions, factors such as the model of the GenAI tool, the content and quality of prompts given to these models, and how the AI processes the input may affect the outcomes.

With the increasing prevalence of GenAI tools and their growing integration into education, questions have arisen regarding whether AI can replace teachers. According to the quantitative results of this study, the feedback provided by the chatbot fulfilled the function of teacher feedback. Considering the time spent by the teacher per student for feedback (12 minutes on average) compared to the time spent by the chatbot (4 minutes on average; see Appendix 9), GenAI tools can play the role of an assistant teacher, reducing teachers' cognitive workload in Turkish writing education. Many studies support our finding regarding AI's potential to reduce teachers' workload (Fan & Ma, 2022; Guo & Wang, 2024; Tossell et al., 2024). As a result, the teacher's predominant role as a source of knowledge may shift toward the affective domain. In short, rather than replacing teachers, GenAI may serve as a transformative technology that enhances teachers' efficiency and speed, modeling an education system where teacher–AI interaction is optimized (Karatas & Yüce, 2024). Furthermore, recent studies indicate that students tend to prefer a combined feedback approach that integrates feedback from teachers, peers, and generative AI tools (Zeevy-Solovey, 2024).

One of the most prominent qualitative findings of this study is that the chatbot was able to provide feedback similar to that of a teacher, coming close to replicating teacher feedback. Indeed, in their study comparing the quality and accuracy of teacher and ChatGPT feedback, Steiss et al. (2024) stated that what was surprising was not that the teacher provided higher-quality feedback than ChatGPT but that ChatGPT was capable of delivering teacher-like feedback. However, analysis of the feedback content in this study revealed that affective expressions were more prominent in teacher feedback. The mnemonic, emphatic, stimulating, and associative informative aspects of teacher feedback, which address the individuality of the student, were lacking in AI-generated feedback, although this may change as AI advances. The teacher autonomy observed in this study, along with the resulting emotional engagement and reactions (such as attitude, motivation, frustration, etc.) of students during the feedback process, may have influenced their writing achievement scores. This finding aligns with several studies (Han & Hyland, 2015; Jansen et al., 2025; Mahfoodh, 2017; Máñez et al., 2024; Renata et al., 2024; Zheng & Yu, 2018).

An analysis of the types of feedback provided on students' texts showed that both the teacher and the GenAI chatbot systematically used directive, informative-stimulating, praise-encouraging, and summary feedback. One of the most critical findings was that while the teacher provided question-based feedback, the GenAI chatbot did not use this type of feedback at all. This suggests that the teacher encouraged students to think critically by asking questions. In their study comparing teacher and ChatGPT feedback, Guo and Wang (2024) found that teachers tend to not only provide informative feedback but also ask questions to stimulate students' thinking. Additionally, in this study, when the percentage distribution of feedback types provided by the chatbot and the teacher was examined, a partially complementary

pattern emerged. For example, AI provided directive feedback 26.82% of the time, while the teacher provided it 14.66% of the time. In contrast, the teacher provided praise feedback 25.86% of the time, compared to 20.86% for AI. This pattern aligns with the findings of Yao et al. (2025), which examined AI-driven chatbot applications for middle school English teachers providing feedback on student writing.

An analysis of feedback focus areas revealed that both teacher and AI feedback primarily emphasized grammar and spelling rules, content, and organization. A notable difference was that the teacher provided fewer but more balanced feedback points, whereas the GenAI chatbot focused more on content. This suggests that teachers provide more specific feedback, taking into account students' previous mistakes and deficiencies, while the AI chatbot offers more general and inclusive feedback. This finding is supported by the results of a study conducted by Castro et al. (2025), which examined the use of CyberScholar—a GenAI assistant designed to provide writing feedback in K–12 education. Furthermore, the smaller quantity of feedback provided by the teacher may be due to the cognitive demands of assessing student writing. However, since this did not negatively affect students' writing achievement, it may suggest that the teacher's feedback was more effective and enduring. This is consistent with Lee's (2019) "less is more" principle, which advocates for focusing on students' specific weaknesses and encouraging critical thinking rather than overwhelming them with excessively detailed corrective feedback.

In conclusion, based on both the quantitative and qualitative findings of this study, GenAI tools may serve as a complementary, innovative, and alternative feedback tool that reduces teachers' cognitive workload during the feedback stage of writing instruction, freeing time for other essential writing processes such as preparation, drafting, planning, and publishing. However, a limitation of using GenAI tools in writing education is that neither teachers nor students currently possess sufficient AI literacy. Therefore, training in AI literacy is recommended for both teachers and students (Tate et al., 2023; van Wyk, 2025).

CONCLUSION, IMPLICATIONS AND SUGGESTIONS

The primary aim of this study was to compare, both quantitatively and qualitatively, the impact of teacher versus GenAI-based chatbot feedback on the development of middle school students' writing skills. The quantitative findings showed no statistically significant difference in students' informative writing performance scores between the AI feedback and teacher feedback groups [$t = -1.428$, $p > 0.05$]. In terms of retention, a comparison of follow-up test scores revealed a decrease in the AI group and a significant increase in the teacher group. Furthermore, beyond the observed difference, the teacher group demonstrated a greater long-term effect compared to the AI group. The qualitative findings revealed a meaningful difference in both feedback types and components of written expression. While the chatbot tended to focus more heavily on content and provided more detailed feedback, the teacher offered more balanced feedback. Additionally, the teacher utilized interrogative feedback—a type not observed in the AI responses.

These findings suggest that GenAI tools hold significant potential for providing feedback (Cope & Kalantzis, 2023; Mollick & Mollick, 2023; Sabzalieva & Valentini, 2023). The first implication is that GenAI-supported tools may be effectively utilized to provide feedback on students' texts in the development of writing skills at the middle school level. These tools can be integrated into the writing process either directly or indirectly. Secondly, the study's retention results indicate that AI tools should not be viewed as alternatives to teachers but rather as innovative and complementary sources of feedback. Thirdly, the fact that the teacher provided less feedback than the AI yet achieved comparable effects in the posttest and a greater effect in the follow-up test may suggest that, despite AI's ability to fulfill the formal and functional aspects of feedback delivery, it is currently unable to replicate the affective dimension of the teacher–student relationship. Moreover, feedback should not be regarded as a simple task of automated response based on predefined criteria but rather as an essential component of the writing process involving affective dimensions (Corbin et al., 2025).

In conclusion, this study addresses a gap in the literature concerning the use of GenAI tools to provide feedback on student writing. However, further research is needed to fully understand both the potential and the limitations of these tools. Future studies could examine how

feedback practices differ between teachers with AI literacy and those without, and how these differences impact students' writing performance. Additionally, future research could compare the accuracy and content of feedback provided by general-purpose GenAI tools (e.g., ChatGPT, Claude, Gemini, DeepSeek) with feedback provided by GenAI chatbots specifically trained for feedback purposes. It would also be valuable to explore the writing development of students who use GenAI tools at home for feedback versus those who rely solely on teacher feedback. Also worth investigating are the effects of teacher and GenAI feedback across different genres of writing and various grade levels. Finally, future research could explore students' perceptions of GenAI feedback and teachers' experiences with these tools.

LIMITATIONS

The findings of this study may not be universally applicable due to the small sample size, the focus on fifth-grade students, the application to teaching informative texts, and the use of a limited dataset to train the AI chatbot. These limitations should be taken into consideration when analyzing the findings of the study.

DATA ACCESSIBILITY STATEMENT

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Appendices.** Appendix 1 to 9. DOI: https://doi.org/10.55982/openpraxis.17.3.871.s1

ETHICS AND CONSENT

The conduct of this research was approved on ethical grounds by the Ethics Committee of Social and Human Sciences at Gaziantep University, based on Decision No. 22 taken at Meeting No. 08 on April 3, 2024. Ethical principles of scientific research were followed at all stages of the study.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS (CRediT)

Mehmet Ali Baz; Conceptualization, investigation, chatbot training, validation, writing—original draft preparation, writing—review and editing. All authors have read and agreed to the published version of the manuscript OR The author has read and agreed to the published version of the manuscript. Sevil Hasırcı Aksoy: Conceptualization, methodology, formal analysis, data analysis and findings, data curation, validation, visualization, writing—review and editing, supervision.

SUSTAINABLE DEVELOPMENT GOALS (SDGs)

This study is linked to the following SDG(s): Quality education (SDG 4).

AUTHOR NOTES

In line with *the Academic Integrity and Transparency in AI-assisted Research and Specification Framework* (Bozkurt, 2024), the authors of this study acknowledge the responsible and transparent use of AI tools at various stages of the research process. During the idea development phase, the basic free versions of Elicit AI and Consensus AI were utilized to support the literature review. The preliminary outputs generated by these tools were critically

evaluated and subsequently refined solely by the human authors in accordance with ethical and academic standards. Potential biases inherent in AI-generated content were also assessed and addressed. All final conceptual frameworks and scholarly arguments are the sole responsibility of the human authors. In the language translation and localization phase, Deepl Pro and ChatGPT Plus (versions as of April 2025) were used to translate the manuscript into English. The translated text was reviewed by an expert in English academic writing instruction, then revised and finalized by the human authors. As the final stage, the manuscript underwent professional proofreading to ensure linguistic precision and academic consistency. For the data analysis, ChatGPT Plus was employed to calculate and verify Cohen's d and r effect sizes. These analyses were reviewed, refined, and finalized by the human authors to ensure accurate representation and reporting of the research data. The authors take full responsibility for the validity and integrity of the reported findings.

AUTHOR AFFILIATIONS

Mehmet Ali Baz orcid.org/0000-0002-6847-5657
Gaziantep University, Türkiye
Sevil Hasırcı Aksoy orcid.org/0000-0003-1041-7558
Gaziantep University, Türkiye

REFERENCES

Abdelhalim, S. M. (2023). Online writing instruction: Understanding university EFL instructors' beliefs and practices as a complex system. *Language Teaching Research*, 13621688221150855. https://doi.org/10.1177/13621688221150855

Akgun, S., & **Greenhow, C.** (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, *2*, 431–440. https://doi.org/10.1007/s43681-021-00096-7

Akkaya, N., & **H. Çıvğın.** (2021). Artificial intelligence in Turkish education. *The Journal of International Education Science*, *8*(29), 308–322. https://doi.org/10.29228/INESJOURNAL.53915

Alsofyani, A. H., & **Barzanji, A. M.** (2025). The Effects of ChatGPT-generated feedback on Saudi EFL learners' writing skills and perception at the tertiary level: A mixed-methods study. *Journal of Educational Computing Research*, *63*(2), 431–463. https://doi.org/10.1177/07356331241307

Angraini, E. P. (2023). *Improving students' descriptive writing ability by using peer assessment at eighth grade of SMP N 2 Sungkai Selatan in 2022/2023 academic year* [Unpublished doctoral dissertation]. UIN Raden Intan Lampung.

Applebee, A. N., & **Langer, J. A.** (2011). EJ Extra: A snapshot of writing instruction in middle schools and high schools. *English Journal*, *100*(6), 14–27. https://doi.org/10.58680/ej201116413

Aridah, A., Atmowardoyo, H., & **Salija, K.** (2017). Teacher practices and students' preferences for written corrective feedback and their implications on writing instruction. *International Journal Of English Lingusitics*, *7*(1), 112–125. http://www.ccsenet.org/journal/index.php/ijel

Ballet, K., & **Kelchtermans, G.** (2009). Struggling with workload: Primary teachers' experience of intensification. *Teaching and Teacher Education*, *25*(8), 1150–1157. https://doi.org/10.1016/j.tate.2009.02.012

Balta, Y., & **Türel, Y. K.** (2013). Çevrimiçi uzaktan eğitimde kullanılan farklı ölçme değerlendirme yaklaşımlarına ilişkin bir inceleme [An analysis of different assessment approaches used in online distance education]. *Electronic Turkish Studies*, *8*(3). https://doi.org/10.7827/TurkishStudies.4271

Bardwell, R. (1981). Feedback: How does it function? *Journal of Experimental Education*, *50*(1), 4–9. https://doi.org/10.1080/00220973.1981.11011792

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, Article 100745. https://doi.org/10.1016/j.asw.2023.100745

Başttürk, R. (2014). Deneme modelleri [Experimental designs]. In A. Tanrıöğren (Ed.), *Bilimsel araştırma yöntemleri* [Scientific research methods] (pp. 31–53). Anı.

Beach, R., & **Friedrich, T.** (2006). Response to writing. *Handbook of writing research*, *1*.

Biber, D., Nekrasova, T., & **Horn, B.** (2011). The effectiveness of feedback for L1-English and L2-writing development: A meta-analysis. *ETS Research Report Series*, *2011*(1), i–99. https://doi.org/10.1002/j.2333-8504.2011.tb02241.x

Bozkurt, A. (2020). Educational technology research patterns in the realm of the digital knowledge age. *Journal of Interactive Media in Education*, *2020*(1), 1–17. https://doi.org/10.5334/jime.570

Bozkurt, A. (2023a). Postdigital artificial intelligence. In P. Jandrić (Ed.), *Encyclopaedia of postdigital science and education*. Springer. https://doi.org/10.1007/978-3-031-35469-4_2-2

Bozkurt, A. (2024). GenAI et al.: Cocreation, Authorship, Ownership, Academic Ethics and Integrity in a Time of Generative AI. *Open Praxis*, *16*(1), 1–10. https://doi.org/10.55982/openpraxis.16.1.654

Burstein, J. (2020). Natural language processing and the literacy challenge. In H. Jiao & R. W. Lissitz (Eds.), *Application of artificial intelligence to assessment* (pp. 77–100). Information Age Publishing, Inc.

Büyüköztürk, Ş. (2007). *Sosyal bilimler için veri analizi el kitabı* (7. b.) [*Data analysis handbook for social sciences* (7th ed.)]. Pegem.

Büyüköztürk, Ş., Çakmak, K. E., Akgün, Ö. E., Karadeniz, Ş., & **Demirel, F.** (2023). Bilimsel araştırma yöntemleri (34. baskı). [Scientific Research Methods (34 th ed.)]. Pegem Akademi.

Castro, V., Nascimento, A. K. D. O., Zheldibayeva, R., Searsmith, D., Saini, A., Cope, B., & **Kalantzis, M.** (2025). Implementation of a Generative AI Assistant in K-12 Education: The CyberScholar Initiative. https://doi.org/10.48550/arXiv.2502.19422

Cen, Y., & **Zheng, Y.** (2024). The motivational aspect of feedback: A meta-analysis on the effect of different feedback practices on L2 learners' writing motivation. *Assessing Writing*, *59*, 100802. https://doi.org/10.1016/j.asw.2023.100802

Çetinkaya, G., & **Hamzadayı, E.** (2015). Yabancı dil olarak Türkçe öğretim sürecinde sözel düzeltme geri bildirimleri: Öğretmen ve öğrenci yeğleyişleri [Oral corrective feedback in teaching Turkish as a foreign language: Teacher and student preferences]. *Turkish Studies*, *10*(3), 285–302. https://doi.org/10.7827/TurkishStudies.7906

Çetinkaya, G., & **Yolcusoy, Ö.** (2020). Okuduğunu anlamayı değerlendirmeye dönük ölçünleştirilmiş görev biçimleri: Yabancılara Türkçe öğretimi ders kitapları üzerine karşılaştırmalı bir inceleme [Standardized task formats for reading comprehension assessment: A comparative analysis on Turkish as a foreign language textbooks]. HAYEF: *Journal of Education*, *17*(2), 176–198. https://doi.org/10.5152/hayef.2020.20021

Chen, Z., Chen, W., Jia, J., & **Le, H.** (2022). Exploring AWE-supported writing process: An activity theory perspective. *Language Learning & Technology*, *26*(2), 129–148. https://doi.org/10125/73482

Cho, K., Schunn, C. D., & **Charney, D.** (2006). Commenting on writing: Typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication*, *23*(3), 260–294. https://doi.org/10.1177/0741088306289261

Chomsky, N., Roberts, I., & **Watumull, J.** (2023, March 8). The false promise of ChatGPT. *The New York Times*. https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Routledge. https://doi.org/10.4324/9780203029053

Cohen, L., Manion, L., & **ve Morrison, K.** (2007). *Research methods in education* (6th ed.). Routledge.

Cope, B., & **Kalantzis, M.** (2023). Creating a different kind of learning management system: TheCGScholar experiment. In M. Montebello (Ed.), *Promoting next-generation learning environments through CGScholar* (pp. 1–18). *IGI Global*. https://doi.org/10.4018/978-1-6684-5124-3.ch001

Corbin, T., Tai, J., & **Flenady, G.** (2025). Understanding the place and value of GenAI feedback: a recognition-based framework. *Assessment & Evaluation in Higher Education*, 1–14. https://doi.org/10.1080/02602938.2025.2459641

Coşkun, E., & **Tamer, M.** (2015). Yazma Eğitiminde Geri Bildirim Türleri ve Kullanımı [Types and Use of Feedback in Writing Education]. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *12*(32), 337–372. https://dergipark.org.tr/tr/download/article-file/183481

Creswell, J. W. (2020). Eğitim araştırmaları: Nicel ve nitel araştırmanın planlanması, yürütülmesi ve değerlendirilmesi (H. Ekşi, Çev. Ed., 3. baskı). [Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research (3th ed.]). EDAM.

Cui, Y., Schunn, C. D., Gai, X., Jiang, Y., & **Wang, Z.** (2021). Effects of trained peer vs. teacher feedback on EFL students' writing performance, self-efficacy, and internalization of motivation. *Frontiers in Psychology*, *12*, Article 788474. https://doi.org/10.3389/fpsyg.2021.788474

Dawson, P., Henderson, M., Ryan, T., Mahoney, P., Boud, D., Phillips, M., & **Molloy, E.** (2023). Technology and feedback design. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, design, and technology: An international compendium of theory, research, practice, and policy* (pp. 695–739). Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-17727-4_124-1

Derakhshan, A., & **Ghiasvand, F.** (2024). Is ChatGPT an evil or an angel for second language education and research? A phenomenographic study of research-active EFL teachers' perceptions. *International Journal of Applied Linguistics*, *34*(4), 1246–1264. https://doi.org/10.1111/ijal.12561

Dibbon, D. (2004). *It's about time. A Report on the Impact of Workload on Teachers and Students*. Faculty of Education, Memorial University of Newfoundland. http://files.nlta.nl.ca/wp-content/uploads/public/documents/wrkldstudy_rprt/wrkldrprt04.pdf

Eckhardt, K. W., & **Ermann, M. D.** (1977). *Social research methods: Perspective, theory, and analysis*. https://lccn.loc.gov/76057256

Ellis, P. D. (2010). The essential guide to effect sizes. Cambridge University Press. https://doi.org/10.1017/CBO9780511761676

Fan, N., & **Ma, Y.** (2022). The Effects of Automated Writing Evaluation (AWE) Feedback on Students' English Writing Quality: A Systematic Literature Review. *Language Teaching Research Quarterly*, *28*, 53–73. https://doi.org/10.32038/ltrq.2022.28.03

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). Sage.

Fleckenstein, J., Liebenow, L. W., & **Meyer, J.** (2023). Automated feedback and writing: A multi-level meta-analysis of effects on students' performance. *Frontiers in Artificial Intelligence*, *6*, Article 1162454. https://doi.org/10.3389/frai.2023.1162454

Flick, U. (1998). *An introduction to qualitative research*. Sage.

George, D., & **Mallery, P.** (2010). *SPSS for Windows step by step: A simple guide and reference*, (17.0 update). Pearson.

Ghani, M., & **Asgher, T.** (2012). Effects of Teacher and Peer Feedback on Students' Writing at Secondary Level. *Journal of Educational Research* (1027–9776), *15*(2).

Gielen, S., Tops, L., Dochy, F., Onghena, P., & **Smeets, S.** (2010). A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum. *British Educational Research Journal*, *36*(1), 143–162. https://doi.org/10.1080/01411920902894070

Graham, S. (2019). Changing how writing is taught. *Review of Research in Education*, *43*(1), 277–303. https://doi.org/10.3102/0091732X18821125

Graham, S., Harris, K., & **Hebert, M. A.** (2011). *Informing writing: The benefits of formative assessment. A Carnegie Corporation Time to Act report*. Alliance for Excellent Education. https://media.carnegie.org/filer_public/37/b8/37b872027138cd6c6f2335bf/ccny_report_2011_informing.pdf?_gl=1*1obn17s*_gcl_au*MTQ4NDIwODI1OS4xNzQ3MTY0MzE4

Graham, S., Hebert, M., & **Harris, K. R.** (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal, 115*(4), 523–547. https://doi.org/10.1086/681947

Graham, S., McKeown, D., Kiuhara, S., & **Harris, K. R.** (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, *104*(4), 879–896. https://doi.org/10.1037/a0029185

Grimes, D., & **Warschauer, M.** (2010). Utility in a Fallible Tool: A Multi-Site Case Study of Automated Writing Evaluation. *Journal of Technology, Learning, and Assessment*, *8*(6). https://files.eric.ed.gov/fulltext/EJ882522.pdf

Guo, K., & **Wang, D.** (2024). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, *29*(7), 8435–8463. https://doi.org/10.1007/s10639-023-12146-0

Gyamfi, G., Hanna, B. E., & **Khosravi, H.** (2024). Impact of an instructional guide and examples on the quality of feedback: insights from a randomised controlled study. *Educational technology research and development*, *72*(3), 1419–1437. https://doi.org/10.1007/s11423-024-10346-0

Hamzadayı, E., & **Dölek, O.** (2022). Bilgilendirici metin yazma başarımını değerlendirmeye yönelik bir ölçek önerisi [A scale proposal for assessing informative text writing performance]. *Gaziantep Üniversitesi Eğitim Bilimleri Dergisi*, *6*(1), 115–130.

Harari, Y. N. (2023, April 28). Yuval Noah Harari argues that AI has hacked the operating system of human civilisation. *The Economist*. https://www.economist.com/by-invitation/2023/04/28/yuval-noah-harari-argues-that-ai-has-hacked-the-operating-system-of-human-civilisation

Han, Y., & **Hyland, F.** (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of second language writing*, *30*, 31–44. https://doi.org/10.1016/j.jslw.2015.08.002

Howell, D. C. (2002). *Statistical methods for psychology* (5th ed.). Duxbury.

Hsieh, Y., Hiew, C. K., & **Tay, Y. X.** (2017). Computer-mediated corrective feedback in Chinese as a second language writing: Learners' perspectives. *Chinese as a second language assessment*, 225–246. https://doi.org/10.1007/978-981-10-4089-4_11

Hyland, K., & **Hyland, F.** (2006). Feedback on second language students' writing. *Language teaching*, *39*(2), 83–101. https://doi.org/10.1017/S0261444806003399

Jansen, T., Höft, L., Bahr, J. L., Kuklick, L., & **Meyer, J.** (2025). Constructive feedback can function as a reward: Students' emotional profiles in reaction to feedback perception mediate associations with task interest. *Learning and Instruction*, *95*, 102030. https://doi.org/10.1016/j.learninstruc.2024.102030

Jerry, C., Jan, J. M., & **Samuel, M.** (2013). Improving students' writing: Why is there a need for teacher feedback. *Malaysian Journal of Languages and Linguistics*, *2*(1), 44–55. https://doi.org/10.24200/mjll.vol2iss1pp44-55

Karataş, F., & **Yüce, E.** (2024). AI and the future of teaching: Preservice teachers' reflections on the use of artificial intelligence in open and distributed learning. *The International Review of Research in Open and Distributed Learning*, *25*(3), 304–325. https://doi.org/10.19173/irrodl.v25i3.7785

Kellogg, R. T., Whiteford, A. P., & **Quinlan, T.** (2010). Does automated feedback help students learn to write? *Journal of Educational Computing Research*, *42*(2), 173–196. https://doi.org/10.2190/EC.42.2.c

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*(1–2), 81–93. https://doi.org/10.2307/2332226

Khan, A. F., Jawad, M., & **Ahmed, Z.** (2025). The influence of ChatGPT-generated feedback on the writing skills of Pakistani ESL high school students: An experimental study. *Research Consortium Archive*, *3*(1). https://rc-archive.com/index.php/Journal/article/download/60/64

Kim, J., Lee, T. H., Bae, Y., & **Kim, M. K.** (2024). A comparison between ai and human evaluation with a focus on generative ai. In Proceedings of the 18th International Conference of the Learning Sciences-ICLS 2024, pp. 1722–1725. *International Society of the Learning Sciences*. https://doi.org/10.22318/icls2024.930382

Kluger, A. N., & **Adler, S.** (1993). Person-versus computer-mediated feedback. *Computers in human behavior*, *9*(1), 1–16. https://doi.org/10.1016/0747-5632(93)90017-M

Kohnke, L., Moorhouse, B. L., & **Zou, D.** (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Lee, E. (2023, March 19). Is ChatGPT a false promise? *Berkeley Blog*. https://blogs.berkeley.edu/2023/03/19/is-chatgpt-a-false-promise

Lee, I. (2019). Teacher written corrective feedback: Less is more. *Language Teaching*, *52*(4), 524–536. https://doi.org/10.1017/S0261444819000247

Liu, Y., Xiong, W., Xiong, Y., & **Wu, Y. F. B.** (2024). Generating timely individualized feedback to support student learning of conceptual knowledge in Writing-To-Learn activities. *Journal of Computers in Education*, *11*(2), 367–399. https://doi.org/10.1007/s40692-023-00261-3

Loncar, M., Schams, W., & **Liang, J. S.** (2023). Multiple technologies, multiple sources: trends and analyses of the literature on technology-mediated feedback for L2 English writing published from 2015–2019. *Computer Assisted Language Learning*, *36*(4), 722–784. https://doi.org/10.1080/09588221.2021.1943452

Lu, Q., Zhu, X., & **Cheong, C. M.** (2021). Understanding the difference between self-feedback and peer feedback: A comparative study of their effects on undergraduate students' writing improvement. *Frontiers in psychology*, *12*, 739962. https://doi.org/10.3389/fpsyg.2021.739962

MacArthur, C. A. (2016). Instruction in evaluation and revision. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 272–287).Guilford.

Mahfoodh, O. H. A. (2017). "I feel disappointed": EFL university students' emotional responses towards teacher written feedback. *Assessing Writing*, *31*, 53–72. https://doi.org/10.1016/j.asw.2016.07.001

Mamad, A., & **Vígh, T.** (2024). EFL/ESL Teachers' and Students' Perceptions and Practices of Written Feedback in a Higher Education Context. *Profile: Issues in Teachers' Professional Development*, *26*(2), 181–198. https://doi.org/10.15446/profile.v26n2.109141

Máñez, I., Lipnevich, A. A., Lopera-Oquendo, C., & **Cerdán, R.** (2024). Examining pre-service teachers' feedback on low- and high-quality written assignments. *Educational Assessment, Evaluation and Accountability*, *36*(2), 225–256. https://doi.org/10.1007/s11092-024-09432-x

Matsumura, L. C., Garnier, H. E., Slater, S. C., & **Boston, M. D.** (2008). Toward measuring instructional interactions "at-scale". *Educational Assessment*, *13*(4), 267–300. https://doi.org/10.1080/10627190802602541

Mayring, P. (2004). Qualitative content analysis. In U. Flick, E. von Kardoff, & I. Steinke (Eds.), *A companion to qualitative research* (pp. 266–269). Sage.

McGee, M., & **Sadler, B.** (2024). Equity in the use of ChatGPT for the classroom: A comparison of the accuracy and precision of ChatGPT 3.5 vs. ChatGPT4 with respect to statistics and data science exams. arXiv preprint arXiv:2412.13116. https://doi.org/10.48550/arXiv.2412.13116

Miles, M. B., & **Huberman, A. M.** (1994). *Qualitative data analysis: An expanded Sourcebook*. (2nd ed). Sage.

Millî Eğitim Bakanlığı [MEB]. (2019). Türkçe dersi öğretim programı (ilkokul ve ortaokul 1-8. sınıflar) [Turkish language curriculum (primary and lower secondary grades 1–8)]. Millî Eğitim Bakanlığı.

Mollick, E., & **Mollick, L.** (2023). *Assigning AI: Seven approaches for students, with prompts* [Preprint]. arXiv. https://doi.org/10.2139/ssrn.4475995

Moore, N. S., & **MacArthur, C. A.** (2016). Student use of automated essay evaluation technology during revision. *Journal of Writing Research*, *8*(1), 149–175. https://doi.org/10.17239/jowr-2016.08.01.05

Nelson, M. M., & **Schunn, C. D.** (2009). The nature of feedback: How diferent types of peer feedback afect writing performance. *Instructional Science*, *37*(4), 375–401. https://doi.org/10.1007/ s11251-008-9053-x

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd Ed.). Sage. https://doi.org/10.1177/1035719X0300300

Peck, L., & **Kavanagh, S. S.** (2024). Learning to give feedback to student writers: A systematic review of research on field-based experiences for pre-service writing teachers. *Teaching and Teacher Education, 144*, 104602. https://doi.org/10.1016/j.tate.2024.104602

Peterson, S., Childs, R., & **Kennedy, K.** (2004). Written feedback and scoring of sixth-grade girls' and boys' narrative and persuasive writing. *Assessing Writing*, *9*(2), 160–180. https://doi.org/10.1016/j.asw.2004.07.002

Pojslová, B. (2022). The role and perception of peer and teacher feedback in multiple-draft feedback provision on foreign language learners' writing. *Pedagogická orientace*, *32*(4), 444–483. https://doi.org/10.5817/PedOr2022-4-444

Raihany, A. (2014). The importance of teacher's written feedback on the students'writing in teaching learning process. okara: *Jurnal Bahasa dan Sastra*, *8*(1), 91–106. https://doi.org/10.19105/ojbs.v8i1.458

Razali, N. M., & **Wah, Y. B.** (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors, and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21–33. https://www.nrc.gov/docs/ml1714/ml17143a100.pdf

Renata, D., Renati, D., Hidayah, N., & **Indriani, L.** (2024). The impact of teacher's feedback on students' writing ability. *English Education:Journal of English Teaching and Research*, *9*(1), 105–113. https://doi.org/10.29407/jetar.v9i1.21669

Rismawati, S. R. A., & **Hikmat, M. H.** (2018). Feedback Given By The Teacher On Students' Writing At The Seventh Grade Of Smp Negeri 2 Juwiring [Unpublished doctoral dissertation] Universitas Muhammadiyah Surakarta. http://eprints.ums.ac.id/id/eprint/61476

Roscoe, R. D., Wilson, J., Johnson, A. C., & **Mayra, C. R.** (2017). Presentation, expectations, and experience: Sources of student perceptions of automated writing evaluation. *Computers in Human Behavior*, *70*, 207–221. https://doi.org/10.1016/j.chb.2016.12.076Get rights and content

Sabzalieva, E., & **Valentini, A.** (2023). *ChatGPT and artificial intelligence in higher education: Quick start guide*. UNESCO.

Saliu-Abdulahi, D., Hellekjær, G. O., & **Hertzberg, F.** (2017). Teachers' (formative) feedback practices in EFL writing classes in Norway. *Journal of Response to Writing*, *3*(1), Article 3. https://scholarsarchive.byu.edu/journalrw/vol3/iss1/3

Sarı, E., & **Han, T.** (2022). An investigation of Turkish EFL instructors' beliefs about writing in English. *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi*, *12*(1), 169–184. https://doi.org/10.48146/odusobiad.1080702

Sarıkaya, H. (2023). Yazılı metinleri ölçme ve değerlendirmede geri bildirim kaynakları [Feedback sources in assessing and evaluating written texts]. *Uluslararası Filoloji Bengü*, *3*(2), 29–46. https://doi.org/10.29228/filolojibengu.31

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL quarterly*, *41*(2), 255–283. https://doi.org/10.1002/j.1545-7249.2007.tb00059.x

Shermis, M. D. (2014). The challenges of emulating human behavior in writing assessment. *Assessing Writing*, *22*, 91–99. https://doi.org/10.1016/j.asw.2014.07.002

Shermis, M. D., & **Hamner, B.** (2013). Contrasting state-of-the-art automated scoring of essays. In M. D. Shermis, & J. Burstein (Eds.), *Handbook of automated essay evaluation: Current applications and new directions* (pp. 313–346). Routledge. https://doi.org/10.4324/9780203122761.ch19

Shi, H., & **Aryadoust, V.** (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, *36*(2), 187–209. https://doi.org/10.1017/S0958344023000265

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Srichanyachon, N .(2012). Teacher written feedback for L2 learners' writing development. *Humanities, Arts and Social Sciences Studies*, 7–17. https://so02.tci-thaijo.org/index.php/hasss/article/view/7149/6174

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & **Olson, C. B.** (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, *91*, 101894. https://doi.org/10.1016/j.learninstruc.2024.101894

Stevenson, M., & **Phakiti, A.** (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, *19*, 51–65. https://doi.org/10.1016/j.asw.2013.11.007

Sullivan, G. M., & **Feinn, R.** (2012). Using effect size or why the p value is not enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1

Tabachnick, B. G., & **Fidell, L. S.** (2013). *Using multivariate statistics*. Pearson.

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

Tate, T., Doroudi, S., Ritchie, D., Xu, Y., & **Warschauer, M.** (2023). Educational research and AI-generated writing: Confronting the coming tsunami. EdArXiv. January, 10. https://www.writecenter.org/uploads/1/2/5/6/125695261/educational_research_and_ai-generated_writing_tate_et_al_2023__2_.pdf

Teddlie, C., & **Yu, F.** (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, *1*(1), 77–100. https://doi.org/10.1177/1558689806292430

Timperley, H. S., & **Parr, J. M.** (2010). What is this lesson about? Instructional processes and student understandings in writing lessons. In C. Rubie-Davies (Ed.), *Educational psychology: Concepts, research and challenges* (pp. 32–49). Routledge.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & **Agyemang, B.** (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 1–24. https://doi.org/10.1186/s40561-023-00237-x

Tossell, C. C., Tenhundfeld, N. L., Momen, A., Cooley, K., & **de Visser, E. J.** (2024). Student perceptions of ChatGPT use in a college essay assignment: Implications for learning, grading, and trust in artificial intelligence. *IEEE Transactions on Learning Technologies, 17*, 1069–1081. https://ieeexplore.ieee.org/document/10400910

Ülper, H. (2019). Yazma becerilerini ölçmede kullanılan farklı ölçme yaklaşımlarının geçerlikleri üzerine, kompozisyon yazmanın amaçları bağlamında karşılaştırmalı bir inceleme [A comparative study on the validity of different assessment approaches used in measuring writing skills in the context of the purposes of composition writing]. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, *52*(1), 1–30. https://www.researchgate.net/publication/332736492

van Wyk, M. M. (2025). Student Teachers' Leveraging GenAI Tools for Academic Writing, Design, and Prompting in an ODeL Course. *Open Praxis, 17*(1), 95–107. https://doi.org/10.55982/openpraxis.17.1.711

Wahyudi, A., Sari, F., Pamuji, A., & **Dharmawan, Y. Y.** (2024). From Critique to Insight: Student Voices on English Writing Feedback. *Voices of English Language Education Society*, *8*(1), 282–292. https://doi.org/10.29408/veles.v8i1.25659

Ware, P. D., & **Warschauer, M.** (2006). Electronic Feedback and Second Language Writing. In K. Hyland, & F. Hyland (Eds.), *Feedback in Second Language Writing* (pp. 105–122). Cambridge University Press. https://doi.org/10.1017/CBO9781139524742.008

Warschauer, M., & **Grimes, D.** (2008). Automated writing assessment in the classroom. *Pedagogies: An International Journal, 3*(1), 22–36. https://doi.org/10.1080/15544800701771580

Wilson, J., Myers, M. C., & **Potter, A.** (2022). Investigating the promise of automated writing evaluation for supporting formative writing assessment at scale. *Assessment in Education: Principles, Policy & Practice*, *29*(2), 183–199. https://doi.org/10.1080/0969594X.2022.2025762

Wilson, J., & **Czik, A.** (2016). Automated essay evaluation software in English Language Arts classrooms: Effects on teacher feedback, student motivation, and writing quality. *Computers & Education, 100,* 94–109. https://doi.org/10.1016/j.compedu.2016.05.004

Wilson, J., & **Roscoe, R. D.** (2020). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research*, *58*(1), 87–125. https://doi.org/10.1177/0735633119830764

Wisniewski, B., Zierer, K., & **Hattie, J.** (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in psychology*, *10*, 487662. https://doi.org/10.3389/fpsyg.2019.03087

Wondim, B. M., Bishaw, K. S., & **Zeleke, Y. T.** (2024). Effectiveness of teachers' direct and indirect written corrective feedback provision strategies on enhancing students' writing achievement: Ethiopian university entrants in focus. *Heliyon, 10*(2). https://doi.org/10.1016/j.heliyon.2024.e24279

Xiao, J., Bozkurt, A., Nichols, M., Pazurek, A., Stracke, C. M., Bai, J. Y. H., Farrow, R., Mulligan, D., Nerantzi, C., Sharma, R. C., Singh, L., Frumin, I., Swindell, A., Honeychurch, S., Bond, M., Dron, J., Moore, S., Leng, J., Slagter van Tryon, P. J., Garcia, M., Terentev, E., Tlili, A., Chiu, T. K. F., Hodges, C. B., Jandrić, P., Sidorkin, A., Crompton, H., Hrastinski, S., Koutropoulos, A., Cukurova, M., Shea, P., Watson, S., Zhang, K., Lee, K., Costello, E., Sharples, M., Vorochkov, A., Alexander, B., Bali, M., Moore, R., Zawacki-Richter, O., Asino, T. I., Huijser, H., Zheng, C., Sani-Bozkurt, S., Duart, J. M., & **Themeli, C.** (2025.) Venturing into the Unknown: Critical Insights into Grey Areas and Pioneering Future Directions in Educational Generative AI Research. *TechTrends,* 1–16. https://link.springer.com/article/10.1007/s11528-025-01060-6

Xie, Y., Huang, L., & **Wang, Y.** (2020). The impact of AWE and peer feedback on Chinese EFL learners' English writing performance. In *Technology in Education. Innovations for Online Teaching and Learning: 5th International Conference, ICTE 2020*, Macau, China, August 19–22, 2020, *Revised Selected Papers 5* (pp. 258–270). Springer, Singapore. https://doi.org/10.1007/978-981-33-4594-2_22

Xiu-Yi, W. U. (2024). AI in L2 learning: a meta-analysis of contextual, instructional, and social-emotional moderators. *System*, 103498. https://doi.org/10.1016/j.system.2024.103498

Yancey, K., et al. (2004). CCCC position statement on teaching, learning, and assessing writing in digital environments. *Coll. Compos. Commun.*, *55*(4), 785–790. https://doi.org/10.58680/ccc20042788

Yang, H., & **Kyun, S.** (2022). The current research trend of artificial intelligence in language learning: a systematic empirical literature review from an activity theory perspective. *Australasian Journal of Educational Technology, 38*(5), 180–210. https://doi.org/10.14742/ajet.7492

Yang, L., & **Li, R.** (2024). ChatGPT for L2 learning: Current status and implications. *System*, *124*, 103351. https://doi.org/10.1016/j.system.2024.103351

Yao, Y., Zhu, X., Xiao, L., & **Lu, Q.** (2025). Secondary school English teachers' application of artificialintelligence-guided chatbot in the provision of feedback on student writing: An activity theory perspective. *Journal of Second Language Writing*, *67*, 101179. https://doi.org/10.1016/j.jslw.2025.101179

Zeevy-Solovey, O. (2024). Comparing peer, ChatGPT, and teacher corrective feedback in EFL writing: Students› perceptions and preferences. *Technology in Language Teaching & Learning*, *6*(3), 1482–1482. https://doi.org/10.29140/tltl.v6n3.1482

Zhai, N., & **Ma, X.** (2022). Automated writing evaluation (AWE) feedback: A systematic investigation of college students' acceptance. *Computer Assisted Language Learning*, *35*(9), 2817–2842. https://doi.org/10.1080/09588221.2021.1897019

Zheng, Y., & **Yu, S.** (2018). Student engagement with teacher written corrective feedback in EFL writing: A case study of Chinese lower-proficiency students. *Assessing Writing*, *37*, 13–24. https://doi.org/10.1016/j.asw.2018.03.001

Zou, D., Luo, S., Xie, H., & **Hwang, G. J.** (2022). A systematic review of research on flipped language classrooms: Theoretical foundations, learning activities, tools, research topics and findings. *Computer Assisted Language Learning*, *35*(8), 1811–1837. https://doi.org/10.1080/09588221.2020.1839502

TO CITE THIS ARTICLE:
Baz, M. A., & Hasirci Aksoy, S. (2025). The Effect of Feedback on Informative Text Writing: AI or Teacher? *Open Praxis*,*17*(3), pp. 611–631. DOI: https://doi.org/10.55982/openpraxis.17.3.871

Submitted: 16 March 2025
Accepted: 06 May 2025
Published: 11 August 2025