International Journal of Educational Development
ISSN Print: 3104-4263
ISSN Online: 3104-4271
CD ISSN: 3104-428X
https://doi.org/10.63313/IJED.9008

A Case Study on the Impact of ChatGPT Feedback on Underachieving Students' Writing Ability in Senior High School English Instruction

Lanqiao Zhou
China West Normal University, Nanchong 637000, China

How to cite this paper: Zhou, L. Q. (2025). A case study on the impact of ChatGPT feedback on underachieving students' writing ability in senior high school English instruction. International Journal of Educational Development, 1(1), 52–59. ISSN Print: 3104-4263, ISSN Online: 3104-4271.
https://doi.org/10.63313/IJED.9008
Published: 2025-07-15

Abstract

This case study investigates the influence of ChatGPT-generated feedback on the writing abilities of an underachieving senior high school student in China. Over a six-week period, the participant engaged in weekly writing tasks, used ChatGPT for iterative revisions, and submitted final drafts for teacher evaluation. Data were collected through writing scores, semi-structured interviews, classroom observations, and student reflective logs. Results indicate measurable gains in grammatical accuracy, organizational coherence, and writing confidence, alongside increased use of varied vocabulary and transitional devices. However, evidence of potential overdependence on AI suggestions highlights the need for structured pedagogical frameworks. The study offers practical recommendations for integrating AI-assisted feedback into English writing instruction to support low-performing learners.

Keywords

ChatGPT; Writing feedback; Case study

1. Introduction

Writing is recognized as one of the most complex language skills, requiring synthesis of vocabulary, grammar, and higher-order thinking such as argumentation and logical structure (Harmer, 2015). In Chinese senior high school settings, underachieving students often struggle with writing due to limited language foundation, low motivation, and insufficient individualized feedback (Wen Qiufang, 2016). Traditional instructional models—focused on rote memorization and formulaic essay structures—frequently fail to address the specific needs of these learners, resulting in persistent low performance (Hyland, 2003).

With the advent of large language models like ChatGPT, educators have gained access to tools capable of providing immediate, personalized feedback on multiple aspects of writing, including grammatical accuracy, lexical variety, and organizational

suggestions (Chen & Zhao, 2023). ChatGPT's 24/7 availability and interactive design allow students to engage in iterative revision cycles, potentially mitigating anxiety and promoting autonomy (Ellis, 2009).

This study focuses on one Grade 11 underachieving student (Student X) whose writing performance consistently fell below average. The research aims to answer three core questions:

(1) How does an underachieving student utilize ChatGPT feedback during English writing tasks?

(2) What specific effects does ChatGPT feedback have on the student's writing accuracy, coherence, and confidence?

(3) Which pedagogical strategies can optimize the integration of AI-generated feedback for learners requiring additional support?

By addressing these questions, the study seeks to contribute both theoretical insights and practical guidelines for AI-assisted writing instruction in high school contexts.

2. Literature Review

2.1. Underachieving Students and Writing Challenges

Underachieving students are defined as those whose academic outcomes fall significantly below expectations despite adequate instructional input (Wang Qiang, 2012). In English writing, they often exhibit:

(1) Limited vocabulary range and frequent lexical inaccuracies (Ferris, 2003).

(2) Grammatical errors in tense, agreement, and article usage due to inadequate meta-cognitive strategies.

(3) Poor organizational coherence, with underdeveloped paragraph unity and weak use of transitional phrases.

(4) Low motivation and high anxiety, which further hinder writing practice and engagement (Dörnyei, 2001).

Traditional feedback mechanisms—teacher written comments and peer reviews—are often insufficient. Teachers, constrained by large class sizes, prioritize error marking over in-depth guidance, while peer feedback quality is variable and sometimes detrimental for low-proficiency learners (Lundstrom & Baker, 2009).

2.2. Theoretical Perspectives on Writing Feedback

Feedback is crucial for writing development and can be categorized by focus (form vs. content) and delivery (direct vs. indirect) (Brookhart, 2017).

Direct feedback supplies explicit corrections, which aids low-level learners but risks dependency. Indirect feedback encourages self-correction, promoting deeper processing but may frustrate underachievers without scaffolding (Ellis, 2009).

Selective feedback—targeting specific error types—can prevent cognitive overload,

particularly for struggling writers (Ferris, 2007). Content-focused comments address ideas, organization, and development, fostering higher-order thinking (Hyland, 2003).

2.3. AI-Assisted Feedback and ChatGPT

Recent studies highlight AI's potential in writing instruction. Automated essay scoring systems like e-rater® provide form-focused feedback efficiently but lack interactive dialog (Attali & Burstein, 2006).

ChatGPT extends these capabilities by allowing multi-turn interaction, offering paraphrasing options, transitional suggestions, and brief outlines, thus supporting both form and superficial content guidance (Chen & Zhao, 2023).

Nevertheless, challenges include model hallucinations, variable depth in content feedback, and risks of the student outsourcing cognitive tasks to AI rather than internalizing learning (Bender et al., 2021).

To harness AI effectively, teachers must integrate prompting strategies, reflection logs, and balanced assessment practices to ensure students remain active authors of their work (Wen Qiufang, 2016).

3. Methodology

This chapter presents the overall methodological design of the case study, covering the research approach, participant selection, data collection instruments, data analysis procedures, and measures taken to ensure ethical considerations and reliability.

3.1. Research Design and Approach

A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident (Stake, 1995). Here, the phenomenon of interest is the impact of ChatGPT-based feedback on an underachieving student's writing progress, and the context is a Chinese senior high school. Given the complexity of classroom teaching, institutional constraints, and individual psychological factors, a single-case design allows in-depth exploration of how ChatGPT feedback interacts with the student's specific circumstances (Merriam, 1998).

The study employs a qualitative-dominant design supplemented by a limited amount of quantitative data (e.g., writing scores). Qualitative data collection methods (e.g., interviews, observations, logs) provide rich insights into the student's evolving attitudes, strategies, and perceptions, which pure quantitative scores cannot fully capture. Meanwhile, writing performance measures across multiple tasks offer a partial indicator of improvement or stagnation in specific dimensions (e.g., grammar, organization).

3.2. Participant Selection and Setting

The study took place in a senior high school, characterized by relatively large class sizes of around 45–50 students per class. The English teachers at this school face significant administrative and teaching loads, commonly leaving them with limited time for extensive feedback on student writing drafts. The institutional environment thus reflects typical challenges found in many Chinese high schools.

The researcher identified a Grade 11 student—pseudonymously named Student X—meeting the following criteria:

(1) Consistent Low Performance: Persistently falling below the class average on English exams, especially in writing, for at least two consecutive semesters.

(2) Language Deficits: Reports from the teacher indicated frequent grammatical inaccuracies, limited vocabulary, and difficulty organizing written text.

(3) Low Motivation and Self-Confidence: According to teacher observation and the student's prior self-reports, Student X displayed a sense of helplessness regarding writing.

(4) Willingness to Participate: Student X voluntarily agreed to be the focal participant, with the consent of his parents/guardians, after being briefed on the study's aims and procedures.

3.3. Data Collection Instruments

To achieve methodological triangulation, the study utilized multiple instruments:Writing Tasks,Teacher and Student Interviews,Classroom Observations,Student Logs,Writing Performance Evaluation (Quantitative Scoring).

1. Writing Tasks

During the six-week intervention, the teacher assigned a 120–150 word English composition each week, spanning various genres (narrative, argumentative, expository). The writing procedure was standardized:

(1) Initial Draft: Student X composed the first draft under near-exam conditions (20–25 minutes).

(2) ChatGPT Feedback: After completing the draft, Student X pasted the text into ChatGPT, requesting corrections or improvements. Several interaction turns were allowed, enabling the student to ask for clarifications or additional suggestions.

(3) Final Draft: Based on ChatGPT's feedback and any subsequent teacher guidance, Student X revised the composition and submitted it for grading.

2. Teacher and Student Interviews

Teacher Interviews: Conducted once before the intervention (for background information) and once at the end (for summative insights). The pre-intervention interview focused on identifying Student X's prior writing performance, learning attitude, and any relevant contextual details. The post-intervention interview explored the teacher's observations on Student X's progress, ChatGPT usage, and broader pedagogical implications.

Student Interviews: Conducted weekly (8–10 minutes each) to investigate how Student X interacted with ChatGPT, what kind of feedback was most or least helpful, and any evolving attitudes toward writing. These short but regular interviews were audio-recorded and transcribed for subsequent analysis.

3. Classroom Observations

The researcher visited Student X's English class twice (Weeks 2 and 5) to observe general writing instruction, teacher feedback patterns, and any explicit references to ChatGPT usage. Field notes were taken to document classroom interactions, the teacher's approach to writing tasks, and how Student X responded or participated.

4. Student Logs

Student X was required to maintain a concise weekly log of approximately 100–150 words, reflecting on the following aspects:

(1) Experiences with using ChatGPT for the week's writing tasks.
(2) Notable improvements achieved or challenges encountered during the process.
(3) Personal emotions related to writing, such as motivation, anxiety, or satisfaction.
(4) Additional observations or questions considered relevant.

These reflections served as a valuable source of qualitative data, providing insights into the student's evolving perceptions and self-reflective practices throughout the study.

5. Writing Performance Evaluation

Each initial and final draft was scored by the teacher based on a 100-point rubric addressing four dimensions:

1. Language Accuracy (grammar, vocabulary) – 40%
2. Content Relevance and Depth – 20%
3. Organization and Cohesion – 20%
4. Overall Fluency – 20%

This system mirrored standard high school English writing rubrics, providing a quantitative measure of writing improvement. By comparing initial and final draft scores over the six weeks, the researcher could assess the degree to which ChatGPT feedback facilitated changes in writing quality.

3.4. Data Analysis

3.4.1. Quantitative Analysis

For each weekly assignment, a spreadsheet recorded both the initial draft score and the final draft score. A simple descriptive analysis (mean, standard deviation, score difference) was performed to identify improvement patterns. Although the case study design does not permit robust statistical generalizations, these numerical indicators support the investigation of potential gains attributable to ChatGPT feedback.

3.4.2. Qualitative Analysis

All interview transcripts, student logs, and field notes underwent a thematic analysis (Braun & Clarke, 2006). The researcher followed these steps:

1. Transcription and Immersion: Repeated reading of all transcripts for a holistic sense of data.
2. Initial Coding: Labeling segments of text (e.g., “grammar correction,” “motivation boost,” “frustration with AI suggestions,” “teacher oversight”).
3. Theme Development: Grouping related codes into broader, recurring themes such as “Accuracy Gains,” “Overreliance on AI,” “Attitude Shift,” or “Instructional Constraints.”
4. Theme Refinement: Cross-checking themes against the data, merging or splitting as needed to ensure they represent distinct ideas.
5. Interpretation: Linking the identified themes back to the research questions, aided by excerpts from interviews and logs as supporting evidence.

4. Results and Discussion

4.1. Quantitative Findings

Final drafts showed consistent improvements of 6–11 points. Grammar error rates dropped by approximately 30% across the intervention period, while use of linking phrases increased notably (e.g., 'Furthermore,' 'Nevertheless'), suggesting enhanced cohesion.

Week	Initial Draft	Final Draft	Difference (Δ)
1	52	58	+6
2	53	60	+7
3	50	61	+11
4	54	63	+9
5	55	65	+10
6	57	66	+9

4.2. Qualitative Findings

ChatGPT Usage Patterns:Student X primarily requested grammar corrections and paraphrases, averaging 1–3 revision cycles per task. He engaged in interactive negotiation, clarifying suggestions to match his intended meaning.

(1) Motivation and Confidence:Reflective logs revealed reduced writing anxiety and increased willingness to revise drafts. The student described ChatGPT as a 'safety net,' fostering greater self-efficacy in subsequent independent writing attempts.

(2) Overreliance Risks:Observations surfaced instances where Student X inserted AI-generated complex phrases without full comprehension, prompting teacher intervention to encourage active reflection.

4.3. Pedagogical Implications

The results and discussion highlight that ChatGPT-based feedback can significantly affect an underachieving student's English writing performance, particularly in the realm of surface accuracy, vocabulary enrichment, and incremental confidence-building. The study underscores both the promise and pitfalls of using AI in writing instruction: while immediate assistance can accelerate improvements, insufficiently regulated or guided AI usage may prompt superficial gains or outright dependence. The role of the teacher remains critical in modeling effective prompting, ensuring ethical standards, and guiding students to internalize, rather than merely adopt, AI feedback. These findings pave the way for the final chapter, which consolidates the major conclusions, addresses study limitations, and proposes practical and theoretical implications for future research and classroom practice.

5. Conclusion

In conclusion, this research highlights the transformative yet complex nature of AI-based tools like ChatGPT in supporting underachieving high school students' writing development. While the immediate improvements in grammatical accuracy and organization are promising, sustainable growth depends on structured teacher mediation, reflective writing practices, and well-crafted guidelines that balance innovation with authenticity. As ChatGPT and analogous technologies advance, the ongoing dialogue among teachers, researchers, and policymakers will shape how effectively such tools are harnessed to nurture confident, capable writers—especially those who stand to benefit the most from timely, individualized feedback in a traditionally time-constrained educational environment.

References

[1] Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater® V.2. The Journal of Technology, Learning, and Assessment, 4(3), 1-30.

[2] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the ACM Conference on Fairness, Accountability, and Transparency.

[3] Brookhart, S. M. (2017). How to Give Effective Feedback to Your Students (2nd ed.). ASCD.

[4] Chen, L., & Zhao, M. (2023). Exploring the Educational Potential of ChatGPT in Second Language Writing. Language Teaching Research, 27(1), 45-59.

[5] Creswell, J. W., & Plano Clark, V. L. (2018). Designing and Conducting Mixed Methods Research (3rd ed.). SAGE.

[6] Dörnyei, Z. (2001). Motivational Strategies in the Language Classroom. Cambridge University Press.

[7] Ellis, R. (2009). Corrective Feedback and Teacher Development. L2 Journal, 1(1), 3-18.

[8] Ferris, D. R. (2003). Response to Student Writing: Implications for Second Language Students. Routledge.

[9] Harmer, J. (2015). The Practice of English Language Teaching (5th ed.). Pearson.

[10] Hyland, K. (2003). Second Language Writing. Cambridge University Press.

[11] Lundstrom, K., & Baker, W. (2009). To Give is Better than to Receive: The Benefits of Peer Review to the Reviewer's Own Writing. Journal of Second Language Writing, 18(1),

30-43.

[12] Stake, R. E. (1995). The Art of Case Study Research. SAGE.

[13] Ware, P. (2014). Technology and Second Language Writing: A Framework for Feedback in the Language Classroom. Language Learning & Technology, 18(2), 16-31.

[14] Wen, Qiufang. (2016). Learning characteristics and teaching strategies for students with low English achievement. Educational Research, 35(2), 45-50.

[15] Wang, Qiang. (2012). A study on the feedback mechanism in English writing teaching. Foreign Languages Teaching and Research, 4, 123-128.