INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 15, Issue 2, (2026) E-ISSN: 2226-6348

# The Impact of Multisource Formative Assessment on Argumentative Writing Development among Chinese EFL University Students

Diao Jianwei, Priscilla Shak*, Daron Benjamin Loo

Centre for the Promotion of Knowledge and Language Learning, Universiti Malaysia Sabah, Malaysia

Email: diao_jianwei_du23@iluv.ums.edu.my, daronloo@ums.edu.my

*Corresponding Author Email: pshak@ums.edu.my

**DOI Link:** http://dx.doi.org/10.6007/IJARPED/v15-i2/28270

***Published Online:*** 28 May 2026

**Abstract**

Although argumentative writing is a crucial skill for English as a Foreign Language (EFL) university students, instruction in many contexts remains largely product-oriented, with limited opportunities for feedback and revision. This study investigated the effectiveness of formative assessment (FA), which included ongoing feedback and revision. A quasi-experimental design was conducted with two intact first-year EFL classes, with one serving as the experimental group and the other as the control group. Both classes completed pre-tests and post-tests, but only the experimental class received multisource FA interventions, including peer assessment, self-assessment, and teacher feedback. Results showed that the experimental group improved significantly in post-test performance compared to pre-test, $t(64) = -6.96$, $p < .001$, with a large effect size ($d = 0.86$), whereas the control group showed only modest improvement, $t(65) = -2.44$, $p = .018$, $d = 0.30$. Between-group comparison confirmed that the experimental class outperformed the control class in post-test scores, $t(129) = -3.63$, $p < .001$, $d = 0.63$. These findings indicate that FA could support argumentative writing development through timely feedback, self-regulation, and collaborative revision. The study offers practical implications for EFL instructors seeking to improve students' writing skills through formative practices.

**Keywords:** Argumentative Writing, English as a Foreign Language, Formative Assessment, Writing Performance, Quasi-Experimental Study

## Introduction

In the academic context, argumentative writing is considered a foundational skill. While its importance has long been emphasized in first language education, it has also attracted increasing scholarly attention in EFL writing research in recent years (Darmawansah et al., 2025; Guo et al., 2022; Nejmaoui, 2019). Argumentative writing is important because it fosters both linguistic and higher-order thinking skills. By requiring students to formulate claims, assess evidence, and address counterarguments, it promotes critical thinking rather than mere opinion expression (Nejmaoui, 2019; Tan, 2023). In addition, argumentative writing strengthens students' reasoning and argumentation skills by requiring them to justify

ideas logically and organize them coherently (Deane, 2018; Hyland, 2019; Ilyas & Arifin, 2025; Ma, 2024). Therefore, argumentative writing is not only a key component of language learning but also an essential tool for developing critical, logical, and independent thinkers. However, despite its recognized importance, argumentative writing is also widely regarded as one of the most challenging tasks for EFL learners because it places considerable demands on their linguistic, cognitive, and rhetorical abilities (Inderawati et al., 2023; Murtadho, 2021; Wale & Bishaw, 2020). Previous research on writing as a complex cognitive and social activity has highlighted the multiple processes involved in planning, generating, and organizing text (Li & Wang, 2024). Furthermore, challenges in argumentative writing may be intensified further by learners' developing linguistic resources and knowledge of disciplinary rhetorical conventions (Bian & Wang, 2016; Hyland, 2019; Rahmat, 2019). Such challenges may become even more prominent in educational contexts where writing instruction emphasizes final products over process-oriented development.

In China, academic writing also holds an important position; nonetheless, focus is often placed upon the quality of final written products over the writing process itself, despite ongoing curriculum reform that calls upon writing instructors and language practitioners to shift the focus on the process of developing language and writing skills (Cheng & Zhang, 2021; Zhang & Cheung, 2018). As a result, students may receive limited process-oriented guidance, including insufficient feedback for meaningful revision and sustained writing development (Guo & Xu, 2021; Yan et al., 2024). Given that argumentative writing involves recursive processes of planning, revising, and reflection, assessment approaches that support ongoing learning and revision may be particularly valuable. In contrast to these summative assessment practices, FA aims to enhance student learning, particularly when it provides detailed feedback on strengths and weaknesses rather than numerical scores alone (Morris et al., 2021).

A meta-analytic review commissioned by the Australian Education Research Organisation (2024) found that FA is beneficial across different student subgroups and subject areas. Through a review of studies published between 2014 and 2024, Song and Mukundan (2025) highlighted a shift in tertiary EFL writing assessment from traditional summative evaluation toward formative, learner-centered approaches. This shift is also evident in Chinese university EFL classrooms. For example, Yan et al. (2024) examined university students' perceptions of multisource FA activities including peer assessment, self-assessment, and teacher feedback, and found that FA promoted self-regulated learning and English proficiency. While their study provided insights regarding FA, their data was qualitative, which could not demonstrate or quantify the effectiveness of FA. Against this backdrop, this study will conduct a quasi-experimental study with pre- and post-tests to examine the effectiveness of the intervention on argumentative writing performance.

The study will be guided by the FA frameworks proposed by Sadler (1989), Black and Wiliam (1998, 2009), and Wiliam and Thompson (2008). Drawing on these frameworks, five pedagogically relevant strategies will be incorporated into classroom practice: clarifying goals, eliciting evidence of learning, providing feedback, peer revision, and self-regulation. These strategies are operationalized as a multisource FA intervention designed to improve Chinese EFL students' argumentative writing performance. This study contributes to the literature on EFL writing and assessment by providing empirical evidence from a quasi-experimental

investigation of FA in a Chinese university EFL context. It examines measurable changes in students' argumentative writing performance through a multisource FA intervention. The findings also extend existing research by demonstrating how FA can function as an integrated instructional approach rather than a set of isolated assessment techniques.

## Literature Review

### *Formative Assessment and Writing Development*

Evidence indicates that FA enhances student achievement across contexts (Lee et al., 2020). In writing instruction, FA helps students identify weaknesses, revise drafts, and improve performance (McNamara & Kendeou, 2022; Palermo & Thomson, 2019). Its positive effects have been reported in relation to writing performance, revision ability, learning attitudes, and learner autonomy (Burner, 2016; Guo & Xu, 2021; Mohamadi, 2018; Prastikawati & Wiyaka, 2020). Previous studies have further demonstrated the positive effects of FA on EFL students' argumentative writing development. For example, Zhong and Yang (2021) found that FA practices involving peer feedback, reflective journals, assessment sheets, and e-portfolios significantly enhanced students' argumentative writing performance and promoted logical thinking, self-regulated learning, and the quality of peer feedback. These practices reflect the integration of multiple assessment sources, including teacher feedback, peer feedback, and self-assessment, within the writing process. Similarly, Burner (2016) emphasized that feedback, revision, self-assessment, and active learner involvement are essential conditions for FA to function effectively in EFL writing classrooms. By integrating multiple perspectives into classroom assessment, multisource FA may provide learners with richer feedback, greater opportunities for self-regulation, and more sustained engagement in argumentative writing tasks.

Recent studies increasingly suggest that FA practices are more likely to support English writing development when they involve explicit standards, learner participation, and iterative revision. Peer assessment can promote improvement when students engage with peers' texts, comment on them using clear criteria, and revise accordingly, with positive effects reported for content, organization, and language use (Shang, 2022; Storch, 2019). Self-assessment can be effectively supported by explicit criteria, such as checklists or rubrics, which help students evaluate their work against defined standards (Andrade et al., 2010; Teng, 2022; Xiao & Yang, 2019). Similarly, criteria-referenced self-assessment has been shown to promote student learning and achievement by helping learners identify strengths and areas for improvement (Andrade et al., 2010). In addition, process-oriented FA involving repeated cycles of drafting, feedback, and revision tends to be more beneficial than assessment focused exclusively on final products, as meaningful writing development often occurs during revision (Burner, 2016; Mohamadi, 2018; Palermo & Thomson, 2019). Research also suggests that combining teacher feedback, peer feedback, and self-reflection may yield stronger outcomes than relying on a single feedback source, because different sources can contribute to different aspects of writing development (Demirel & Enginarlar, 2016; Teng, 2022; Xiao & Yang, 2019).

Despite growing recognition of the potential of FA to support writing development, several important gaps remain in the literature. In particular, classroom-based evidence from experimental, quasi-experimental, and longitudinal studies remains limited, making it difficult to determine how and under what conditions FA is effective (Lam, 2015; Weber et al., 2024).

In addition, the mechanisms through which FA facilitates writing development remain insufficiently specified (Fiskerstrand & Gamlem, 2023; Vasu et al., 2020; Zhong & Yang, 2021). While prior studies have pointed to such factors as feedback, goal setting, peer assessment, and reflection, how these elements interact to produce improvement in writing has not yet been systematically examined (Fiskerstrand & Gamlem, 2023; Vasu et al., 2020; Walls & Johnston, 2020; Weber et al., 2024; Zhong & Yang, 2021).

In response to these research gaps, the present study investigates the impact of FA practices on students' argumentative writing performance in a Chinese university EFL context. Conceptualizing FA as an instructional intervention, the study examines how integrated multisource FA practices—including teacher feedback, peer-supported revision, and self-assessment—can contribute to writing improvement over time. By adopting a quasi-experimental design, this study seeks to provide theoretically grounded and pedagogically relevant evidence regarding the role of FA in enhancing argumentative writing outcomes in higher education.

*Formative Assessment as an Instructional Intervention*

FA is viewed as a process through which evidence of learning is elicited, interpreted, and used to support improvement. A central assumption is that feedback can promote learning only when students understand the target, identify the gap between current and desired performance, and act on that information (Sadler, 1989). In classroom settings, this process is commonly associated with explicit criteria, feedback, learner participation, and opportunities for revision (Black & Wiliam, 1998, 2009; Wiliam & Thompson, 2008). Recent studies suggest that these elements are most effective when implemented systematically as an integrated instructional approach rather than as isolated techniques (Lee, 2025). On this basis, the present study conceptualizes FA as a criteria-guided, feedback-oriented, and learner-engaged intervention in argumentative writing instruction. More specifically, multisource FA practices are understood as integrated classroom enactments that combine teacher-led actions, such as clarifying writing goals, eliciting evidence of student understanding, and providing process-focused feedback, with learner-engaged practices, such as peer revision and self-regulation. The framework assumes that when these elements are systematically integrated into classroom instruction, they can support improvement in students' argumentative writing performance. This conceptualization is illustrated in Figure 1.

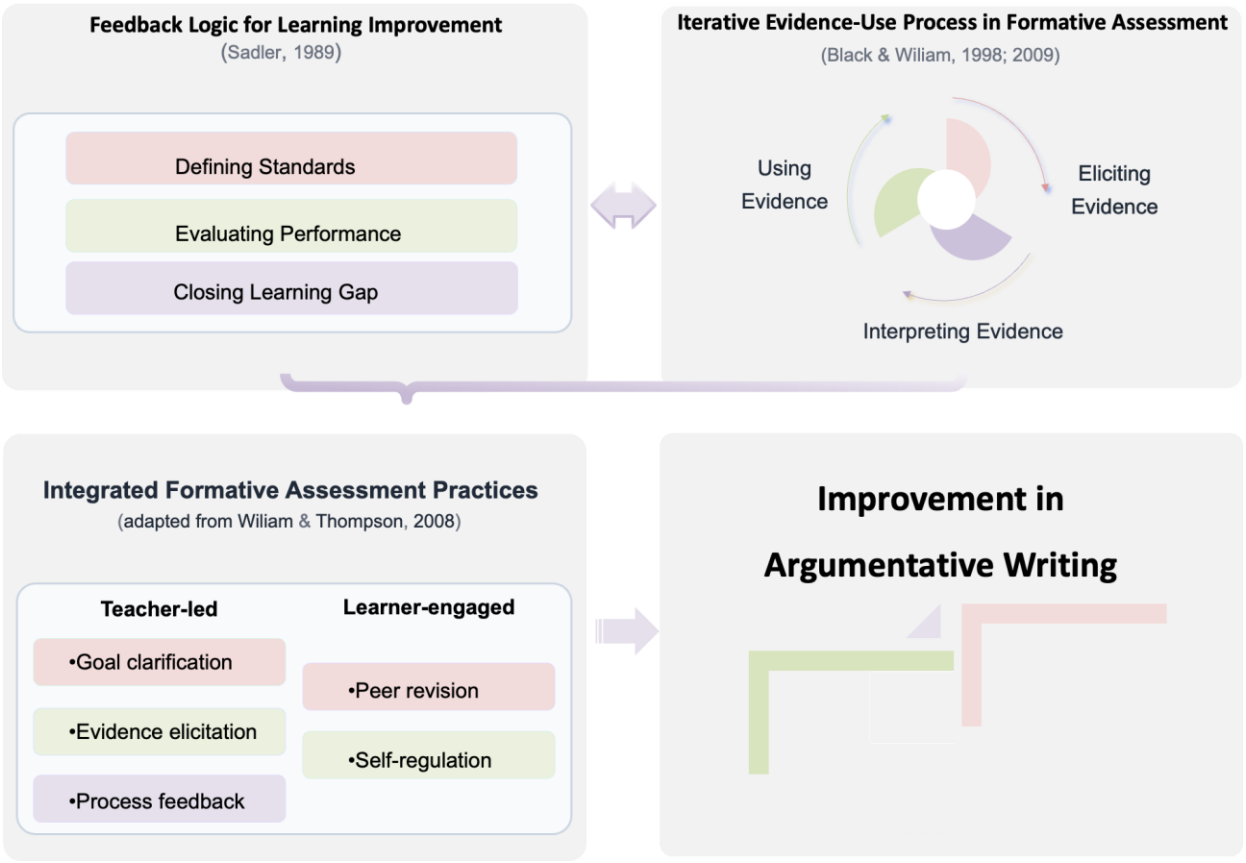


Figure 1 *Conceptual Framework of Formative Assessment and Writing Improvement*

*Research Questions and Hypotheses*

The present study examines the effect of FA on EFL university students' argumentative writing. The investigation was guided by the following research questions and corresponding hypotheses:

RQ1: Does the FA intervention improve the experimental group's argumentative writing performance?

$H_{a1}$: The FA intervention will significantly improve the experimental group's argumentative writing performance.

RQ2: Is the FA intervention more effective than traditional instruction in improving argumentative writing performance?

$H_{a2}$: The FA intervention will be more effective than traditional instruction in improving argumentative writing performance.

## Methodology

A quasi-experimental pre-test–post-test design was adopted in this study to examine the impact of FA practices on students' argumentative writing performance in a university EFL context. As illustrated in Figure 2, the research procedure involved participant grouping, instructional intervention, scoring procedures, and statistical analysis.

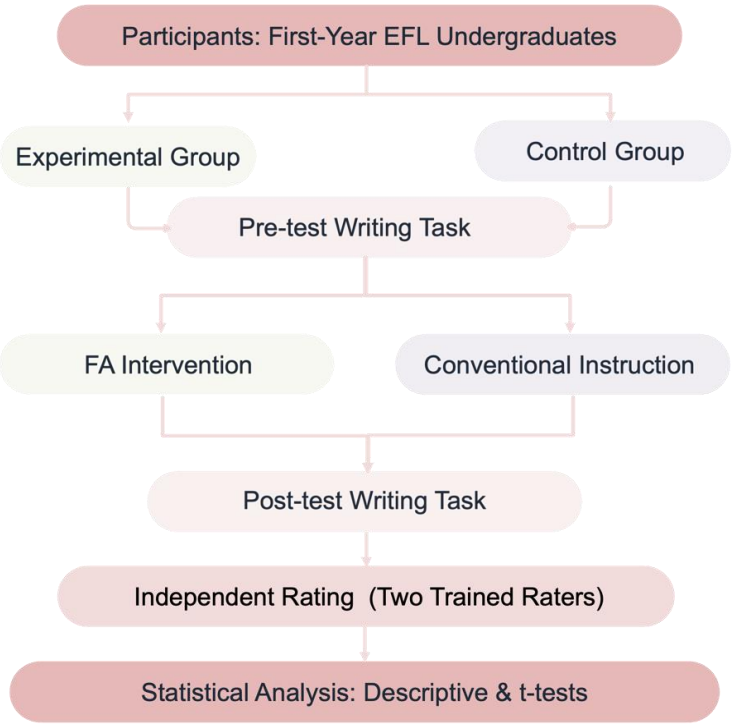


Figure 2 *Quasi-Experimental Research Design of the Writing Intervention Study*

*Participants and Context*

The participants were first-year undergraduate EFL students drawn from two intact classes at a comprehensive university in eastern China. The two classes, each consisting of approximately 70 students, were designated as the experimental group and the control group. Such class sizes are typical of College English instruction at the university, where large-class teaching is common. Using intact classes also allowed the FA intervention to be implemented under regular instructional conditions. All participants had completed the National College Entrance Examination and shared relatively similar English learning backgrounds and proficiency levels, as reflected in their comparable entrance examination scores and pre-test writing performance. The use of intact classes with similar class size and consistent instructional conditions supported the comparability of the two groups prior to the intervention.

Both classes shared the same instructor, curriculum, and teaching materials. The compulsory English writing course followed the university's regular alternating weekly schedule, with four 45-minute sessions in one week and two 45-minute sessions in the next throughout the semester. The instructor was also the researcher who implemented the intervention and had over ten years of teaching experience. Although this dual role may have introduced potential researcher bias and expectancy effects, standardized instructional procedures and analytic scoring rubrics were applied throughout the study to reduce these threats. In addition, having the same instructor teach both groups helped ensure consistency in instructional delivery. The study was conducted within the regular classroom setting without altering the existing institutional timetable or course structure, thereby preserving the authenticity of the instructional environment.

*Design and Intervention*

At the beginning of the semester, both groups completed an argumentative writing pre-test under comparable conditions. During the 16-week semester, the experimental group participated in two structured FA writing cycles initiated in weeks 4 and 9, respectively. In each cycle, the teacher clarified the writing goals and assessment criteria, students produced an initial draft, engaged in peer review using a checklist on content, language, and argumentative strategies, revised their drafts in response to peer comments, and then completed a self-assessment task based on a rubric and written comments. Further revisions were made by the students where necessary before teacher evaluation. In contrast, the control group followed conventional writing instruction, which mainly emphasized summative evaluation and offered limited structured opportunities for feedback-based revision. At the end of the semester, both groups completed a post-test writing task.

*Measures and Data Analysis*

Students' argumentative writing performance was measured using pre-test and post-test tasks adapted from previous CET-4 writing prompts, which are widely used in China to assess the English proficiency of non-English-major university students. The tasks were selected to ensure an appropriate level of difficulty and relevance to university EFL learners. Students were required to compose essays of at least 120 words within 30 minutes. Writing performance was evaluated using a holistic scoring rubric consistent with standard CET-4 writing assessment practices. The rubric assesses overall writing quality by integrating content development, organizational clarity, language use, and overall coherence into a single score. All essays were independently rated by two experienced university English teachers, each with more than six years of experience in evaluating CET-4 writing sections. The use of parallel writing tasks and standardized scoring procedures helped enhance comparability and reliability in performance measurement.

Prior to conducting parametric tests, assumptions of normality and homogeneity of variances were examined. Normality was assessed using the Shapiro–Wilk test, which indicated that the data did not significantly deviate from a normal distribution ($p > .05$ for all groups). Homogeneity of variances was confirmed by Levene's test ($p > .05$). Given the robustness of t-tests to moderate violations of normality and the confirmation of baseline equivalence, parametric tests were deemed appropriate.

For data analysis, descriptive statistics were first calculated to summarize the mean scores and standard deviations of both groups. Paired-sample t-tests were conducted to examine within-group changes from pre-test to post-test, while independent-sample t-tests were used to compare differences between the experimental and control groups. Effect sizes (Cohen's d) were calculated to evaluate the magnitude of the observed differences. These analyses enabled the study to determine the extent to which FA practices contributed to improvements in students' argumentative writing performance.

*Validity, Reliability, and Ethics*

Validity was addressed through the selection of writing tasks. Both pre-test and post-test argumentative writing tasks were drawn from previous CET-4 writing prompts, standardized assessments developed by language testing experts (Jin & Jie, 2024). A holistic scoring rubric aligned with CET-4 practices was employed to evaluate overall writing quality. Two

experienced CET-4 raters independently scored all essays after rubric training. To strengthen internal validity in this quasi-experimental design, the experimental and control groups were intact classes with identical instructional conditions (same instructor, curriculum, and materials). Baseline equivalence was confirmed by pre-test scores, the intervention spanned a full semester to mitigate history effects, and consistent testing procedures were applied across groups (Shadish et al., 2002). Reliability was established through standardized scoring procedures and independent double rating. Inter-rater reliability was examined using the intraclass correlation coefficient (ICC = 0.81), indicating good agreement between the two raters (Koo & Li, 2016). Discrepancies were resolved through discussion, and parallel writing tasks were used for pre-test and post-test to ensure measurement stability.

This study strictly adhered to the ethical guidelines for educational research and received ethical approval from the relevant institutional review board prior to data collection. All participants were fully informed of the research purpose and procedures and provided written informed consent before participation. Participants were assured of their right to withdraw at any time without affecting their academic standing. All data were anonymized, used solely for this research, and kept strictly confidential. The intervention was integrated into the regular instruction of the experimental group and posed no additional risks to participants. To ensure equitable benefits, students in the control group received the same instructional intervention after the completion of the study.

**Results**

To examine the impact of FA intervention on students' argumentative writing performance, an independent-samples t-test was conducted to compare the pretest scores between the experimental and control groups. Prior to the t-test, assumptions of normality and homogeneity of variances were checked. Graphical inspection of the Normal Q–Q plots (Figures 3 and 4) showed that the data points for both groups generally followed the diagonal line, indicating approximate normality. This observation was further supported by non-significant Shapiro–Wilk test results (control: p = .132; experimental: p = .157). Levene's test confirmed homogeneity of variances (F = 0.138, p = 0.711). The t-test revealed no significant difference between the experimental group (M = 74.22, SD = 6.47) and the control group (M = 75.00, SD = 6.08), t(129) = -0.45, p = .65. This confirms baseline equivalence and supports the validity of subsequent comparisons.

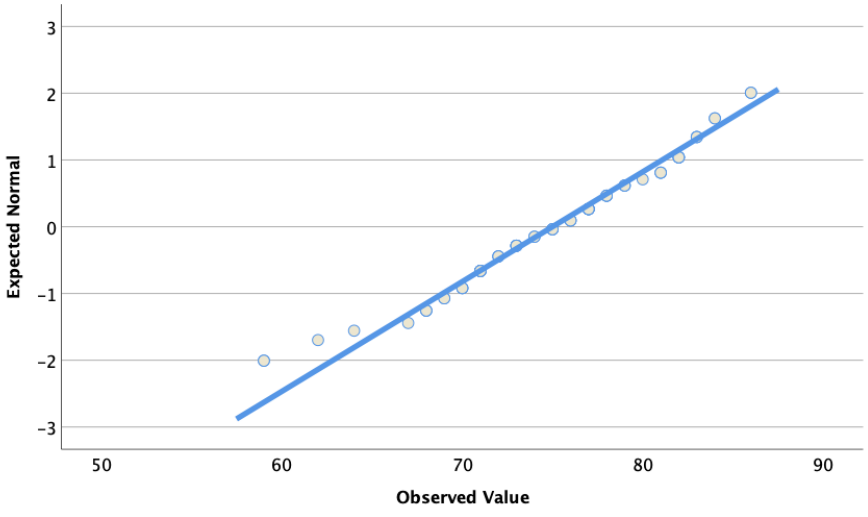


Figure 3 *Normal Q-Q Plot of Pretest Scores for the Control Group*

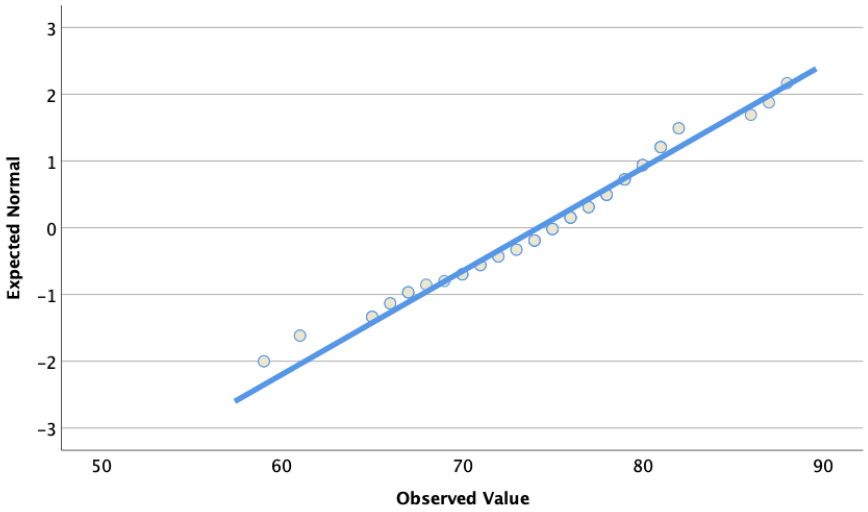


Figure 4 *Normal Q-Q Plot of Pretest Scores for the Experimental Group*

*Within-Group Changes*

Paired-samples t-tests were conducted to examine within-group changes from pre-test to post-test. The control group showed an increase under traditional instruction. Their mean scores rose from 75.00 (SD = 6.08) to 77.59 (SD = 5.55), and although the change was statistically significant, t(65) = -2.44, p = .018, the effect size was relatively small (d = 0.30), indicating limited improvement. The experimental group also showed improvement in writing performance following the multisource FA intervention. Scores increased from pre-test (M = 74.22, SD = 6.47) to post-test (M = 81.02, SD = 5.25), with a statistically significant difference, t(64) = -6.96, p < .001, and a large effect size (d = 0.86). The detailed results for both groups are summarized in Table 1.

Table 1
*Within-Group Changes in Writing Scores*

| Group (N) | Pre-test (M ± SD) | Post-test (M ± SD) | t(df) | p | Cohen's d |
|---|---|---|---|---|---|
| Exp (65) | 74.22 ± 6.47 | 81.02 ± 5.25 | -6.96 (64) | < .001 | 0.86 |
| Ctrl (66) | 75.00 ± 6.08 | 77.59 ± 5.55 | -2.44 (65) | .018 | 0.30 |

*Note*. Exp = experimental group; Ctrl = control group; M = mean; SD = standard deviation.

*Between-Group Comparison*

To examine between-group differences at post-test, an independent-samples t-test was conducted. The experimental group (M = 81.02, SD = 5.25) achieved significantly higher scores than the control group (M = 77.59, SD = 5.55), t(129) = -3.63, p < .001, with a medium-to-large effect size (d = 0.63). These results indicate that students exposed to FA outperformed those receiving traditional instruction. The between-group comparison is presented in Table 2.

Table 2
*Between-Group Comparison of Post-test Scores*

| Group (N) | Post-test (M ± SD) | t(df) | p | Cohen's d |
|---|---|---|---|---|
| Exp (65) | 81.02 ± 5.25 | -3.63 (129) | < .001 | 0.63 |
| Ctrl (66) | 77.59 ± 5.55 | – | – | – |

*Note*. Exp = experimental group; Ctrl = control group; M = mean; SD = standard deviation.

*Distributional Patterns of Writing Score Changes*

In addition to mean score improvements confirmed by t-tests, boxplots revealed how score distributions changed across groups and testing occasions. Both groups showed improvement from pre-test to post-test. The boxplots (Figure 5) indicate a more pronounced upward shift in the experimental group. Notably, the entire box moved into a higher score region at post-test, with the median increasing from approximately 75 to 81 and the interquartile range shifting from about 70–79 to 77–85. The upward movement of the lower edge of the box suggests improvement among lower-performing students, while the simultaneous rise in the upper quartile reflects gains among higher-performing students.

In contrast, changes in the control group were relatively modest, with only a relatively small increase in the median and little change in the position or spread of the box. Although the median rose slightly from around 75 to 78, the position and height of the box remained broadly similar across testing occasions, indicating relatively stable score dispersion. Individual observations shown on the boxplots further support this pattern. At post-test, experimental-group scores appear more densely clustered within the 80–88 range, whereas control-group scores remain more widely scattered, with several observations still located in lower performance bands.

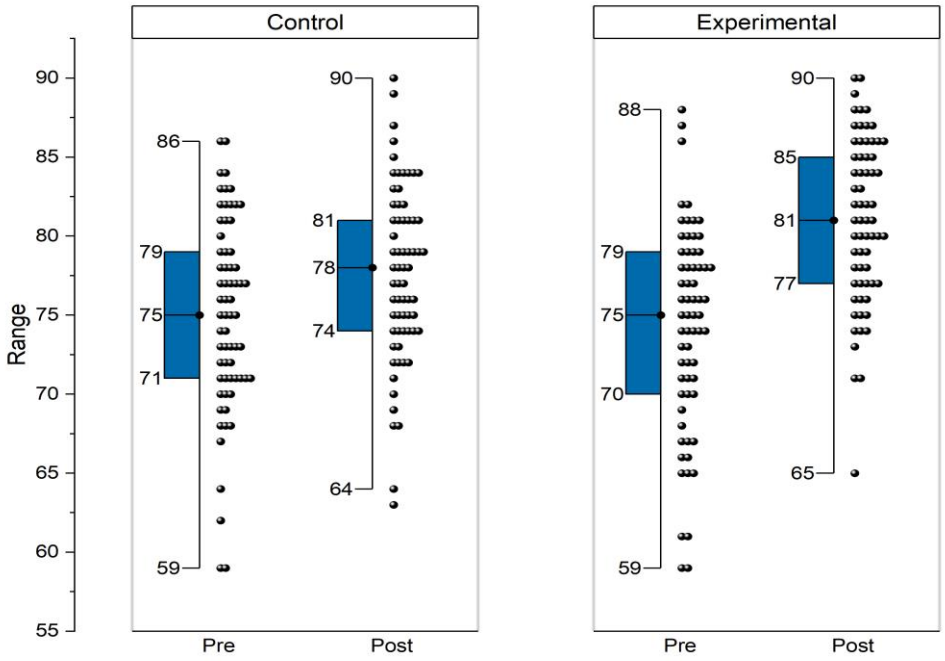


Figure 5 *Distribution of Writing Scores by Group and Testing Occasion*

In combination, these visual patterns suggest that the intervention was associated with an upward redistribution of scores across the performance range rather than a substantial reduction in score variability. Visual inspection of score distributions provides converging descriptive evidence for the statistically significant improvement observed in the experimental group.

**Discussion**

This study found that the experimental group demonstrated significantly greater improvement in argumentative writing performance than the control group. Although both

groups made progress, the gain in the experimental group was substantially larger. These results indicate that multisource FA could effectively enhance writing development in the university context.

*Interpretation of the Intervention Effects*

The effectiveness of FA observed in this study may be accounted for by the theoretical frameworks of the study. Effective feedback requires three conditions: defining standards, evaluating performance, and closing the gap (Sadler, 1989). In the present study, teachers clarified writing standards using rubrics; students evaluated their own performance through ongoing feedback and peer assessment; and subsequent revisions were aimed at reducing the discrepancy between current performance and target standards. Meanwhile, evidence of student writing was elicited through classroom observations, written drafts, peer feedback, and self-assessment. Teachers and students jointly interpreted this evidence, and students subsequently used it to guide their revisions. This iterative process enabled continuous improvement for writing (Black & Wiliam, 1998, 2009). At the level of concrete strategies, the integrated framework proposed by Wiliam and Thompson (2008) illustrates how FA may be operationalized in classroom practice. Teachers set goals, elicited evidence, and provided process-oriented feedback, while students engaged in peer revision and self-regulation. The synergy between these two dimensions likely contributed to the observed gains.

Both groups made progress, but the experimental group improved more. Two reasons may account for the control group's progress. First, regular instruction itself may have had some effect. Second, practice effects from taking the pre-test may have contributed to better post-test performance (Shadish et al., 2002). However, the control group's progress was relatively limited. Previous theoretical and empirical work on FA may help explain this pattern. According to Sadler (1989), the control group may have experienced only the first two steps of effective feedback, i.e., defining standards and evaluating current performance without fully engaging in the third condition of closing the gap. In accordance with Black and Wiliam (1998, 2009), the control group may have remained largely at the stage of eliciting evidence (e.g., receiving a score) without fully interpreting and using that evidence to improve performance. From the perspective of Wiliam and Thompson (2008), the control group lacked the learner-engaged strategies, such as peer revision and self-regulation, which may support the internalization and use of feedback. In contrast, the experimental group engaged more fully in the feedback cycle: they clarified standards, evaluated their performance, and closed the gap through targeted revision. They participated in the iterative process of eliciting, interpreting, and using evidence to guide their writing. In addition, the combination of teacher-led support and learner-engaged practices may have strengthened the effectiveness of the intervention. These theory-informed interpretations help explain why the experimental group showed greater improvement.

As first-year non-English-major undergraduates, many participants may have had limited prior experience with peer assessment, self-assessment, and iterative revision practices. In this context, the use of multisource feedback may have reduced ambiguity regarding writing expectations and provided students with clearer performance standards and more opportunities for reflection and revision. Rather than relying solely on teacher correction, students in the experimental group actively participated in peer review, self-assessment, and iterative revision, which may have contributed to their greater improvement in

argumentative writing performance. The intervention was implemented within a large-class, examination-oriented Chinese EFL context, where writing instruction has traditionally been characterized by teacher-centered practices and a strong emphasis on final written products and test performance. In such settings, opportunities for sustained formative interaction and learner-centered revision practices are often limited. Therefore, the integration of multisource formative feedback in the present study may be considered particularly relevant, as it attempted to introduce process-oriented writing development within a traditionally product-oriented instructional environment. This suggests that FA approaches, when implemented through multisource feedback, may be adaptable to traditional EFL classroom environments. The novelty of the present study lies in its implementation as a multisource feedback system integrating teacher feedback, peer assessment, and self-assessment within a unified criteria-based framework.

The effectiveness of FA may also be closely related to the cognitive and structural demands of argumentative writing, which requires students to construct logical arguments, organize ideas coherently, and support claims with evidence. These genre-specific demands make argumentative writing particularly responsive to feedback and revision processes. In the present study, the multisource FA approach provided students with repeated opportunities to refine their arguments, reorganize their ideas, and improve coherence through iterative feedback cycles. This sustained engagement with feedback and revision may help explain the observed improvement in students' overall writing performance.

The findings of the present study have implications for theory, practice, and pedagogy. Although FA is not a new concept and its positive effects on writing performance have been widely documented (e.g., Cheng & Zhang, 2021; Guo & Xu, 2021; Yan et al., 2024), the present study suggests that its effectiveness may be better understood in terms of how it is implemented rather than whether it works. Specifically, the findings indicate that the integration of multisource feedback may strengthen the FA cycle by providing clearer performance standards and multiple opportunities for revision. From a theoretical perspective, this implies that FA should not be conceptualized as a single feedback mechanism, but rather as an interactive system involving multiple sources of assessment (teacher, peer, and self).

The study also has practical implications for curriculum design and assessment. The positive effects observed in the experimental group suggest that structured multisource FA practices, such as teacher feedback, peer review, and self-assessment, can be meaningfully incorporated into regular writing instruction to support students' development over time. In this respect, the study indicates that writing instruction may benefit from moving beyond a predominantly summative orientation toward a more process-focused model that embeds feedback and revision into the curriculum. The use of CET-4-based writing tasks in the present study further suggests that standardized assessment materials can also serve formative purposes when they are integrated with feedback and revision cycles. At the same time, the implementation of FA may require institutional support, including sufficient time for feedback and revision and classroom conditions that enable sustained teacher and student engagement.

Pedagogically, the findings highlight the importance of organizing writing instruction around explicit standards, multiple sources of evidence, and opportunities for revision. Teachers need to make writing expectations visible before students begin writing so that learners can judge their own work against clear criteria. They also need to gather evidence from varied sources, including student drafts, peer feedback, self-assessment, and classroom interaction, in order to develop a fuller understanding of students' progress. Most importantly, feedback should not remain at the level of comments or scores; it needs to be used to guide concrete revision. Structured peer review and self-assessment can play an important role in this process by helping students internalize criteria, reflect on weaknesses, and develop greater self-regulation in writing. In this sense, the pedagogical value of FA lies not in adding more assessment activities, but in ensuring that assessment evidence is used to support improvement.

**Limitations and Future Directions**

Several limitations of the study should be acknowledged, along with directions for future research. One limitation concerns the research design. This study employed a quasi-experimental design with intact classes, as random assignment of students to different classes was not feasible in authentic educational setting. Where circumstances permit, future research employing randomized controlled designs would provide more evidence.

A further limitation concerns the intervention duration. The intervention spanned one semester, leaving it unclear whether the observed effects can be sustained over longer periods. Longitudinal designs across multiple semesters or academic years would help determine the durability of FA effects on writing development.

The study is also limited in terms of outcome measures, which were confined to writing scores. Incorporating process data such as drafts, reflections, or interviews would yield a more comprehensive understanding. It mainly relied on quantitative measures and did not explore in depth students' qualitative experiences with FA. Future research could incorporate qualitative methods, such as interviews or think-aloud protocols, to illuminate how students interpret feedback, engage in self-regulation, and develop metacognitive strategies during FA.

**Conclusion**

This study examined whether FA could improve argumentative writing among university EFL students, and the findings indicate that it can. The experimental group showed significantly greater improvement than the control group, suggesting that FA can effectively support argumentative writing development in this context.

The contribution of the present study lies not merely in examining FA in a university setting, but in showing how major perspectives on FA can be operationalized within a coherent classroom-based intervention. By integrating feedback, evidence use, peer review, self-assessment, and revision into a connected instructional process, the study provides empirical support for FA as an integrated pedagogical approach rather than a set of isolated techniques. It also shows that such an approach can be implemented through authentic regular classroom practices. In this respect, the study contributes not only context-specific evidence, but also a practical model for enacting FA in university writing instruction. Overall, these findings

suggest that FA has considerable value for university EFL writing instruction and warrants further investigation across other contexts and writing tasks.

**References**

Andrade, H. L., Du, Y., & Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education: Principles, Policy & Practice, 17*(2), 199–214. https://doi.org/10.1080/09695941003696172

Australian Education Research Organisation. (2024). *The impact of context on evidence-based practices: A rapid literature scan on formative assessment, explicit instruction and mastery learning*. https://www.edresearch.edu.au/ research/research-reports/impact-context-evidence-based-practices-rapid-literature-scan

Bian, X., & Wang, X. (2016). Chinese EFL undergraduates' academic writing: Rhetorical difficulties and suggestions. *Indonesian Journal of Applied Linguistics, 6*(1), 20–29.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/ 0969595980050102

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Burner, T. (2016). Formative assessment of writing in English as a foreign language. *Scandinavian Journal of Educational Research, 60*(6), 626–648. https://doi.org/10.1080/00313831.2015.1066430

Cheng, X., & Zhang, L. J. (2021). Sustaining university English as a foreign language learners' writing performance through provision of comprehensive written corrective feedback. *Sustainability, 13*(15), 8192.

Darmawansah, D., Rachman, D., Febiyani, F., & Hwang, G.-J. (2025). ChatGPT-supported collaborative argumentation: Integrating collaboration script and argument mapping to enhance EFL students' argumentation skills. *Education and Information Technologies, 30*(3), 3803–3827. https://doi.org/ 10.1007/s10639-024-12800-5

Deane, P. (2018). The challenges of writing in school: Conceptualizing writing development within a sociocognitive framework. *Educational Psychologist, 53*(4), 280–300.

Demirel, E. T., & Enginarlar, H. (2016). Effects of combined peer-teacher feedback on second language writing development. *Hacettepe University Journal of Education, 31*(4), 657–675. https://doi.org/10.16986/HUJE.2016015701

Fiskerstrand, P., & Gamlem, S. M. (2023). Instructional feedback to support self-regulated writing in primary school. *Frontiers in Education*. https://doi.org/10.3389/feduc.2023.1232529

Guo, K., Wang, J., & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing, 54*, 100666. https://doi.org/ 10.1016/j.asw.2022.100666

Guo, Q., & Xu, Y. (2021). Formative assessment use in university EFL writing instruction: A survey report from China. *Asia Pacific Journal of Education, 41*(2), 221–237. https://doi.org/10.1080/02188791.2020.1798737

Hyland, K. (2019). *Second language writing*. Cambridge University Press.

Ilyas, H. P., & Arifin, S. (2025). Critical thinking in EFL students' argumentative writing: Manifestations and challenges. *Voices of English Language Education Society (VELES), 9*(2). https://doi.org/10.29408/veles.v9i2.29656

Inderawati, R., Hayati, R., Marlina, R., Novarita, N., Awalludin, A., & Anam, S. (2023). Argumentative essay and vocabulary enrichment of English students by utilizing Google Translate. *English Community Journal, 6*(2), 131–141.

Jin, Y., & Jie, W. (2024). Assessing writing in China's College English Test Band 4: Validity arguments for task design and rating scales. *Language Testing, 41*(2), 312–328. https://doi.org/10.1177/02655322231204567

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lam, R. (2015). Convergence and divergence of process and portfolio approaches to L2 writing instruction: Issues and implications. *RELC Journal, 46*, 293–308. https://doi.org/10.1177/0033688215597119

Lee, E. (2025). An integrative research review on formative assessment practices in South Korean EFL contexts. *English Teaching, 80*(2), 103–133. https://doi.org/10.15858/engtea.80.2.202506.103

Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K–12 education: A systematic review. *Applied Measurement in Education, 33*(2), 124–140.

Li, J., & Wang, J. (2024). A measure of EFL argumentative writing cognitive load: Scale development and validation. *Journal of Second Language Writing, 63*, 101095.

Ma, L. (2024). Research on the effective ways to improve students' argument construction ability. *Curriculum and Teaching Methodology, 7*, 42–47. https://doi.org/10.23977/curtm.2024.070308

McNamara, D. S., & Kendeou, P. (2022). The early automated writing evaluation (eAWE) framework. *Assessment in Education: Principles, Policy & Practice, 29*(2), 150–182.

Mohamadi, Z. (2018). Comparative effect of online summative and formative assessment on EFL student writing ability. *Studies in Educational Evaluation, 59*, 29–40. https://doi.org/10.1016/j.stueduc.2018.02.003

Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education, 9*(3), e3292.

Murtadho, F. (2021). Metacognitive and critical thinking practices in developing EFL students' argumentative writing skills. *Indonesian Journal of Applied Linguistics, 10*(3), 656–666.

Nejmaoui, N. (2019). Improving EFL learners' critical thinking skills in argumentative writing. *English Language Teaching, 12*(1), 98–109. https://doi.org/10.5539/ elt.v12n1p98

Palermo, C., & Thomson, M. M. (2019). Classroom applications of automated writing evaluation: A qualitative examination of automated feedback. In B. H. Khan (Ed.), *Educational technology and the new world of persistent learning* (pp. 145–175). IGI Global Scientific Publishing.

Prastikawati, E. F., & Wiyaka, W. (2020). Online backchannel as a formative assessment in improving writing skills. *Journal on English as a Foreign Language, 10*(2), 359–384. https://doi.org/10.23971/jefl.v10i2.2044

Rahmat, N. H. (2019). Problems with rhetorical problems among academic writers. *American Journal of Social Sciences and Humanities, 4*(4), 506–515. https://doi.org/10.20448/801.44.506.515

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144. https://doi.org/10.1007/BF00117714

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

Shang, H.-F. (2022). Exploring online peer feedback and automated corrective feedback on EFL writing performance. *Interactive Learning Environments, 30*(4), 4–16. https://doi.org/10.1080/10494820.2019.1629601

Song, Y., & Mukundan, J. (2025). The influence of humanistic education on tertiary English teachers' writing assessment practices: A systematic review. *Frontiers in Education, 10*, Article 1605368. https://doi.org/10.3389/feduc.2025. 1605368

Storch, N. (2019). Collaborative writing as peer feedback. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (2nd ed., pp. 143–162). Cambridge University Press.

Tan, S. (2023). Second language students' critical thinking performance in argumentative writing. *International Journal of Linguistics, Literature and Translation, 6*(10), 112–130. https://doi.org/10.32996/ijllt.2023.6.10.15

Teng, L. S. (2022). Explicit strategy-based instruction in L2 writing contexts: A perspective of self-regulated learning and formative assessment. *Assessing Writing, 53*, 100645. https://doi.org/10.1016/j.asw.2022.100645

Vasu, K., Fung, Y. M., Nimehchisalem, V., & Rashid, S. M. (2020). Self-regulated learning development in undergraduate ESL writing classrooms: Teacher feedback versus self-assessment. *RELC Journal, 53*, 612–626. https://doi.org/10.1177/0033688220957782

Wale, B. D., & Bishaw, K. S. (2020). Effects of using inquiry-based learning on EFL students' critical thinking skills. *Asian-Pacific Journal of Second and Foreign Language Education, 5*, 9. https://doi.org/10.1186/s40862-020-00090-2

Walls, H., & Johnston, M. (2020). The Fast Feedback method: A quasi-experimental study of the use of formative assessment for primary students' writing. *Australian Journal of Learning Difficulties, 26*, 21–46. https://doi.org/ 10.1080/19404158.2020.1862880

Weber, F., Wambsganss, T., & Söllner, M. (2024). Enhancing legal writing skills: The impact of formative feedback in a hybrid intelligence learning environment. *British Journal of Educational Technology, 56*, 650–677. https://doi.org/ 10.1111/bjet.13529

Wiliam, D., & Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Routledge. https://doi.org/ 10.4324/9781315086545

Xiao, Y., & Yang, M. (2019). Formative assessment and self-regulated learning: How formative assessment supports students' self-regulation in English language learning. *System, 81*, 39–49. https://doi.org/10.1016/j.system.2019.01.004

Yan, Q., Zhang, L. J., Cheng, X., & Zhou, S. (2024). Exploring Chinese university EFL students' perceptions of formative assessment: A qualitative study. *System, 125*, 103391. https://doi.org/10.1016/j.system.2024.103391

Zhang, W., & Cheung, Y. L. (2018). Researching innovations in English language writing instruction: A state-of-the-art review. *Journal of Language Teaching and Research, 9*(1), 80–89.

Zhong, Y., & Yang, M. (2021). Formative assessment in higher education classrooms: Second language writing learning. *International Journal of TESOL Studies, 3*(4), 61–78. https://doi.org/10.46451/ijts.2021.12.05