

Exploring dialogic feedback in EAL academic writing: Performance outcomes and student perspectives

Anna Krulatz

Norwegian University of Science & Technology, Norway
anna.m.krulatz@ntnu.no

Georgios Neokleous

Norwegian University of Science & Technology, Norway
georgios.neokleous@ntnu.no

Tülay Dixon

Oxford College of Emory University, USA
tulay.orucu.dixon@emory.edu

Daniel Dixon

Georgia State University, USA
ddixon49@gsu.edu

Abstract

This study examines the effectiveness of dialogic feedback (DF) as an approach to feedback on English as an additional language (EAL) academic writing at the tertiary level, in comparison with traditional written commentaries. Adopting a mixed-methods design, the study investigates both writing performance outcomes and student perceptions of DF. The findings indicate that while DF offers some advantages in supporting aspects of writing development and promoting more consistent performance across learners, its overall impact on writing quality appears to be modest. Qualitative results further suggest that students generally perceive DF as a valuable, though demanding, process that fosters greater reflection, awareness, and engagement with their writing. Taken together, the findings point to DF as a promising, interaction-oriented approach to feedback that supports learner involvement, while also highlighting the need for further research into its implementation and effectiveness.

Keywords: *Academic writing, dialogic feedback, EAL, tertiary education*

1. Introduction

Research suggests that the development of writing skills in English as an additional language (EAL) can be facilitated through process writing and individualized corrective feedback (Ferris, 1997, 2010). For this reason, techniques such as peer-review, one-on-one teacher-student consultations, voice-recorded commentaries, and written commentaries have been employed (Bitchener et al., 2005; Sugita, 2006). Among written techniques, grammar error correction is often used to help students improve the overall quality of writing, with the goal being an increased ability to self-edit. Teachers can either identify the error and provide the correct form (i.e., direct or explicit corrective feedback) or simply indicate that an error exists without providing the correct form (i.e., indirect or

implicit corrective feedback) (Ene & Kosobucki, 2016). Another common approach is using different types of commentaries such as statements, imperatives, questions, criticism, or personal attribution expressed with hedging devices (Hyland & Hyland, 2001). Finally, teachers can provide feedback using rubrics, which contain lists of criteria or expectations, and provide standards for attaining them. Studies promoting the use of rubrics argued that a primary benefit of rubrics lies in the transparency of the assessment process, as rubrics make the criteria and expectations explicit to the students (e.g., Panadero & Jonsson, 2013).

Although the type of feedback provided by teachers undoubtedly plays a role in the development of EAL writing skills, active student involvement in the assessment process needs to be underscored (Dawson et al., 2019; Muncie, 2000). Weigle (2002) stressed the social and dynamic ability of writing by articulating that “writing in the real world is frequently accomplished with the aid of interactions with other people” (p. 177). However, when feedback on writing is provided by the teacher via written commentaries, the teacher typically “[occupies] the [role] of ‘expert’ and ‘evaluator’ [which] means that the learners have a much-reduced level of choice in deciding whether or not to use that feedback” (Muncie, 2000, p. 47). In the last few decades, in response to critiques of the monologic nature of feedback, a shift has occurred in the positioning of feedback from unidirectional information provided by teachers to understandings of feedback as a process in which learners are expected to take active charge of their own further learning (Dawson et al., 2019; Evans, 2013; Molloy & Boud, 2013). As a result, feedback no longer constitutes “an act of information-giving to students” (Boud & Soler, 2016, p. 403). Instead, greater emphasis is placed on the co-productive nature of feedback and its reflexive dialogic engagement.

Dialogic feedback (DF), defined as an interactive approach to feedback and assessment that embraces and encourages learner responsiveness, requires the skill of writing to be perceived as a malleable process where learning and assessment are inextricably blended in a unified activity with the participation and assistance of a more skilled person (e.g., the teacher) who serves as a mediator (Steen-Utheim & Wittek, 2017). In DF, students assume an active role as they interpret, generate, and respond to ideas in an effort to self-regulate the process of learning (Carless et al., 2011; Hounsell, 2007; Winstone et al., 2017). Research suggests that encouraging students to assume agency in the learning process through practices such as a dialog with others and self-assessment leads to enhanced performance on the task at hand and supports the development of independent learning strategies and higher order thinking skills (Andrade et al., 2009). In addition, practices that promote reflexivity and collaborative dialogue foster sustainable learning as they prepare learners for what might be expected from them in the future to operate efficiently in a complex society (Ajjawi & Boud, 2018; Boud & Soler, 2016).

This paper examines the impact of DF, a non-traditional form of feedback, on EAL students’ academic writing at the tertiary level. More specifically, it explores whether differences exist between a group of students who only received written commentaries and those who engaged in a DF task, along with their attitudes towards DF.

2. Literature Review

2.1 Written Corrective Feedback

Written feedback is catered to individual students to help them make effective revisions when improving their written work. To make the process efficient, teachers can use different commentary types for different purposes, including statements, imperatives, and questions but also through praise, criticism, or personal attribution, which could be mitigated using hedging devices (Hyland & Hyland, 2001). Research on the use of written feedback suggests that while it does seem to help students in their revision process, certain types of comments may be more effective in bringing

about successful revision than others (Ferris, 1997; Nurmukhamedov & Kim, 2010; Sugita, 2006). For instance, Nurmukhamedov and Kim's (2010) study revealed that "imperative ... comments produced more substantive and/or effective revisions compared to questions or statement comments" (p. 280). Similarly, Sugita (2006) found the use of imperatives superior to questions or statements. On the other hand, Nurmukhamedov and Kim (2010) concluded that while imperatives could be a functional way to address errors in form, they were less effective when used to address content. This is possibly because imperatives suggest immediate solutions that can be easily applied to grammatical errors, whereas revising content is a more complex and extensive process. Overall, however, it is important to caution that task-level feedback can facilitate appropriate revisions, such changes do not in themselves indicate learning or development (Hattie & Timperley, 2007).

One specific approach to giving feedback on academic writing is written corrective feedback, also known as grammar error correction or written error correction (Truscott, 1996, 1999). This approach aims to help students improve their writing, with the objective being an increased ability to self-edit writing. The teacher can give written corrective feedback in two ways. One is to identify the error and provide the correct form, also referred to as direct or explicit corrective feedback. The other is to simply indicate that an error exists but without providing the correct form, also referred to as indirect or implicit corrective feedback (Ene & Kosobucki, 2016; Ferris & Roberts, 2006). Research regarding the effectiveness of direct and indirect written corrective feedback is inconclusive (Bitchener, 2008; Ene & Kosobucki, 2016; Truscott, 2007). For instance, regarding the effectiveness of written feedback to promote student self-editing, Ferris and Roberts (2006) argued that there seemed to be no immediate advantage to more explicit feedback, while Erlam et al. (2013) found that explicit correction was less effective in enabling learners to self-correct. However, in their meta-analysis, Biber et al. (2011) argued that providing a more detailed written comment can be more effective than simply identifying an error.

2.2 Rubrics

Another approach often used by teachers to give feedback on written texts is rubrics. Several definitions of a rubric exist in the literature, but in the educational context it is generally understood as a document which lists the criteria, or expectations, of a specific assignment, and provides the standards for attaining those, usually from excellent to poor (Andrade et al., 2009; Jonsson & Svingby, 2007).

Research regarding the extent to which the use of rubrics can promote student learning is inconclusive (Andrade et al., 2009; Jonsson & Svingby, 2007; Panadero & Jonsson, 2013). However, studies promoting the use of rubrics argued that the main potential of rubrics lies in the transparency of the assessment process, as rubrics make the criteria and expectations explicit to the students (Jonsson & Svingby, 2007; Panadero & Jonsson, 2013). Panadero and Jonsson (2013) pointed out that rubrics might not only aid the feedback process, but also lessen student anxiety, improve student self-efficacy, and support student self-regulation, "all of which may indirectly facilitate improved student performance" (p. 140).

2.3 Dialogic Feedback

Traditional approaches to written feedback described above do not provide opportunities for self-regulated learning. As a response to this shortcoming, DF has gained recognition as a valuable approach, particularly in contexts where student engagement with feedback is central to learning. While earlier conceptions of feedback in higher education often treated feedback as a one-way transmission of information, more recent research emphasizes its interactive and dialogic nature

(Jwa, 2025). In this perspective, feedback is not merely a product delivered to passive recipients; rather, it is a dynamic process in which learners actively participate in interpreting, questioning, and applying guidance to their writing (Winstone & Carless, 2019). Within this broader framework, written dialogic feedback represents a particularly prevalent asynchronous modality, allowing learners to engage more deeply and reflectively with feedback across iterative writing processes (Jwa, 2025; Merkel, 2018). Despite its growing prominence, research on DF in EAL academic writing remains limited and continues to develop in both scope and methodological diversity.

DF is grounded in socio-constructivist principles, which view learning as socially mediated and co-constructed through interaction (Vygotsky, 1978). DF typically involves iterative exchanges between teachers and students, during which interpretations are shared, meanings are negotiated, and expectations are clarified (Steen-Utheim & Wittek, 2017). Through these interactions, students move beyond simply receiving comments to actively engaging with feedback, questioning suggestions, and integrating insights into their writing practices. The DF process fosters deeper cognitive engagement and supports the development of self-regulatory skills, enabling learners to monitor, evaluate, and adjust their own strategies (Jwa, 2025).

A growing body of research highlights the pedagogical value of dialogic engagement in additional-language writing. DF can enhance traditional written feedback by providing opportunities for clarification, negotiation, and collaborative meaning-making (Hyland & Hyland, 2006; Perpignan, 2003). Such interactions help students better understand their strengths and weaknesses, cultivate autonomy, and critically reflect on their writing. Moreover, DF addresses common limitations of written-only approaches, such as misinterpretation or lack of applicability to future tasks, by allowing feedback to be immediately adapted to students' individual needs. At the same time, written dialogic feedback is not without challenges. For instance, the asynchronous nature of written feedback may constrain opportunities for real-time negotiation of meaning and interaction (Henderikx et al., 2025), while learners may misinterpret comments when opportunities for clarification are limited (Hyland & Hyland, 2006; Walker, 2009).

Empirical studies suggest that DF positively influences multiple dimensions of student learning (Ajjawi & Boud, 2018; Li & Qian, 2024; Steen-Utheim & Wittek, 2017; Susanti et al., 2025; Winstone et al., 2017). Dialogic approaches have been associated with increased affective, behavioral, and cognitive engagement, encouraging learners to assume agentive roles such as asking questions, responding to feedback, and evaluating revisions (Jwa, 2025). In addition, by co-constructing knowledge with teachers, students move from passive recipients to active contributors. This co-authoring of meaning has been linked to greater ownership of texts, enhanced motivation, and improved transfer of learning across tasks.

Research also emphasizes the relational and contextual aspects of DF. Effective dialogic feedback depends on creating a supportive environment in which students feel comfortable expressing ideas and uncertainties (Steen-Utheim & Wittek, 2017). While power asymmetries between teachers and students may remain, with teachers typically guiding the dialogue, these interactions still provide meaningful opportunities for reflection, participation, and development (Carless et al., 2011; Steen-Utheim & Wittek, 2017). Engaging students in dialogue about feedback helps them develop a clearer understanding of disciplinary standards and expectations, which is essential for improving academic writing competence. Importantly, the effectiveness of written dialogic feedback is also contingent on teachers' ability to provide clear, constructive, and balanced comments, highlighting the role of feedback literacy in successful implementation (Carless & Winstone, 2023; Winstone et al., 2017).

Although studies specifically investigating DF in EAL academic writing are still emerging, existing findings are encouraging. Research indicates that dialogic approaches can enhance engagement with feedback, strengthen writing skills, and build confidence in academic writing (Carless & Boud,

2018; Chinpakdee, 2025; Merkel, 2018; Susanti et al., 2025). Students often report positive perceptions of DF, particularly valuing the individualized support and interactive opportunities it provides (Tam, 2021). These findings suggest that DF has the potential to overcome some of the limitations of traditional feedback methods, fostering active learning, sustained engagement, and long-term development of writing abilities (Carless & Boud, 2018; Deng & Sitthitikul, 2025; Merkel, 2018; Yang & Carless, 2013).

Students' beliefs about feedback play an important role in how they engage with dialogic processes. Research shows that learners generally perceive feedback as valuable for improving writing and often prefer clear and explicit guidance (Fernández-Toro & Furnborough, 2018), while positive attitudes to feedback and active engagement with it are associated with the development of writing competencies (Turner, 2023). At the same time, students do not always understand or effectively use feedback, particularly when comments lack clarity or explanation (Walker, 2009). This suggests a potential tension between expectations for directive input and the active engagement required in dialogic feedback. While many students value opportunities for clarification and individualized support, others may be less accustomed to participatory roles and require support in engaging dialogically (Winstone & Carless, 2019). Understanding these beliefs is therefore essential for evaluating how students perceive and benefit from dialogic feedback in academic writing contexts.

The current study extends this line of research by examining the effectiveness of dialogic feedback in EAL academic writing through both qualitative and quantitative measures. It specifically investigates whether the interactive, effort-intensive nature of DF leads to measurable improvements in writing performance and how students perceive its role in supporting their learning.

2.3 Research Questions

The study addresses the following research questions:

- 1 Are there differences in essay scores between EAL students who received written commentaries as compared with EAL students who participated in a DF task?
- 2 How do the students perceive the role of dialogic feedback in supporting their learning?

3 Methodology

3.1 Context and Participants

Data for this study were collected in an English-medium program (EMI) at a major public university in Norway. The participants were 52 sophomore students enrolled in an EAL teacher education program. Nine of the participants identified as males and 43 identified as females. The students were enrolled in two classes (referred to as Dialogic Group and Traditional Group) taught by the same instructor, a specialist in TESOL and applied linguistics with over ten years of teaching experience in EMI contexts. Both groups were taught using the same content, teaching materials, and methods, and they were required to write an academic paper as one of the course assignments. Two drafts of the assignment were required: a first draft, followed by a round of feedback, and then a second and final draft.

The Dialogic Group, consisting of four males and 17 females, received feedback on the first draft of their paper in form of DF. This entailed indirect comments from the instructor provided in

Microsoft Word to draw learners' attention to errors without overtly correcting the error. The students were required to engage in a dialogic interaction with the instructor. Specifically, they were instructed to track their revisions using the track changes feature provided in Microsoft Word and respond to each comment by explaining how they addressed it. A sample of the interactional revision process is illustrated in Examples 1 and 2 below.

Example 1

Draft 1: Learning English through meaningful, non-linguistic content means learning the English language by using it in what may be called "simulated everyday situations".
Listening to, reading, writing and speaking about a topic in a foreign language forces the students to formulate their own sentences and use their language skills to make themselves understood.
Teacher comment: This is a rather broad definition that could also be applied to communicative language teaching. Could you try to make it more specific?
Student response: I agree and have attempted to specify.
Draft 2: Learning English through meaningful, non-linguistic content means learning the English language by using it through a different subject than English. CBLT means to use specific language objectives to learn a second language through content, defined as subject matter.

Example 2

Draft 1: Students will make brochures about indigenous people. They will work in pairs, and each pair will be given two sheets of A4 paper, two pencils, and crayons.
Teacher comment: This is a good task, but what is the purpose of making the brochure? Will the students then decide which of the indigenous people they want to visit? Or are they social activists for this group and defend its rights?
Student response: I have written the purpose of making the brochure in a new paragraph, where I have included a new activity in the last stage of the lesson plan.
Draft 2: The pairs are going to present their brochure for the rest of their group. They have to be creative and present it in an interesting way to make the rest of the group want to buy their brochure.

The Traditional Group, consisting of five males and 26 females, received written corrective feedback including both direct and indirect comments. These students received instructions to track their revisions, but they did not engage in any dialogic exchanges with the instructor.

3.2 Data Collection and Analysis

The study used a mixed-methods design to examine both the impact of feedback type on writing development and students' experiences with DF. Quantitative data from essay ratings enabled a comparison of writing performance between the two groups, providing insight into potential differences in learning outcomes. Complementing this, qualitative data from student reflections offered a nuanced understanding of how learners interpret and engage with the feedback process. Together, these methods allowed for a comprehensive account of both observable outcomes and learner perspectives.

The first set of data consisted of rater scores of the essays. Both drafts were assessed by two raters using a list of 9 specific criteria including content, style, grammar, citation, cohesion and coherence, organization, thesis statement, introduction, and conclusion, all of which were assessed using a 5-point Likert scale: 0 - unacceptable, 1 - weak, 2 - adequate, 3 - good, and 4 - outstanding (see Appendix A). The raters, one male and one female, were experienced instructors of college-level academic writing. Each rater participated in an individual norming session with one of the researchers. The raters scored all drafts before and after treatment (first draft and final draft) and

met two more times to jointly adjust scores that had a 2-point difference (e.g., a score of two versus four). For all the quantitative analyses, an average of the two scores was used.

The writing performance of both groups was compared using Draft 2 scores, including the overall score and scores for each section of the rubric included in Appendix A. Before comparing the two groups on the basis of Draft 2 scores, we used Draft 1 scores to determine whether the two groups could be considered equal in terms of their writing performance before the treatment. Except for one section of the rubric (i.e., organization), group equivalence could be assumed, as indicated by the Mann-Whitney U-tests presented in Table 1, which were statistically non-significant ($p > .05$) and accompanied by small effect sizes (r). Effect size interpretations were informed by the benchmarks provided by Plonsky and Oswald (2014) for L2 research: values around .25 are considered to be small, .40 medium, and .60 large. These small effect sizes were also accompanied by 95% confidence intervals crossing zero, meaning that, in the larger population of learners to which the sample in this study could be generalized, zero is a plausible value for the magnitude of group differences.

Table 1 Descriptive and inferential statistics comparing the Draft 1 scores of two groups

Aspect of writing	Dialogic feedback <i>n</i> = 21		Traditional Feedback <i>n</i> = 31		Dialogic vs. Traditional		
	<i>M</i> (<i>SD</i>)	<i>Md</i> <i>n</i>	<i>M</i> (<i>SD</i>)	<i>Mdn</i>	<i>W</i>	<i>p</i>	<i>r</i> [95% <i>CI</i>]
Overall Score	22.19 (3.99)	23	21.29 (2.32)	21	415.5	.094	.23 [-.06, .49]
Introduction	2.57 (.36)	2.5	2.45 (.47)	2.5	367.5	.409	-.07 [-.36, .22]
Thesis	2.24 (.52)	2	1.84 (.84)	2	409.5	.103	.23 [-.06, .48]
Overview	2.41 (.46)	2.5	2.50 (.50)	2.5	287.5	.460	-.10 [-.37, .16]
Organization	2.55 (.61)	2.5	2.32 (.48)	2.5	438	*.026	.31 [.04, .59]
Markers	2.57 (.53)	2.5	2.48 (.42)	2.5	378	.304	.14 [-.13, .42]
Conclusion	2.12 (.77)	2	1.84 (.62)	2	396	.160	.20 [-.11, .45]
Grammar	2.64 (.62)	3	2.81 (.40)	3	303	.636	-.07 [-.34, .24]
Style	2.67 (.51)	3	2.81 (.40)	3	291	.469	-.10 [-.37, .14]
Citations	2.43 (.55)	2.5	2.24 (.56)	2.5	381.5	.279	.15 [-.12, .42]

*Significant at $p < .05$

For all inferential statistics, Mann-Whitney U-tests were used because the assumption of normality required for independent samples *t*-tests could not be met for both groups. This assumption was checked using histograms, Q-Q plots, and the normality test of Shapiro-Wilk, a test considered to

be more appropriate for smaller samples than a Kolmogorov-Smirnov test (Loewen & Plonsky, 2016). Results from all Shapiro-Wilk tests can be found in Appendix B. An alpha level of .05 was used for all statistical tests. All analyses were conducted in R (R Core Team, 2026).

The second set of data consisted of student responses to an open-ended, anonymous question answered by students in the Dialogic Group at the end of the semester: “What do you think about the requirement regarding text assignment revisions (i.e., a list of changes you made, tracking changes you made, and responding to the instructor’s comments)? Was it helpful or not when revising your paper?” The purpose of the task was to promote reflective thinking skills and learner autonomy through initiating dialogue about learning and to examine learner attitudes to DF. Responses to the evaluation question were coded as positive (e.g., “The list of changes made me see what was missing and what I could do”), negative (e.g., “I think it was a little unnecessary”), and neutral (e.g., “It really made no difference for me”). In the cases when a response contained both a positive and a negative comment, these were counted as separate.

4. Findings

4.1 Overview

This section presents the results of the study in relation to the two main areas of analysis. First, quantitative findings are reported to examine differences in writing performance between the two groups, including overall scores and specific rubric criteria. This is followed by qualitative findings that explore students’ perceptions of dialogic feedback based on their reflective responses. Together, these results provide insight into both learning outcomes and learner experiences.

4.2 Effectiveness of Dialogic Feedback

Figure 1 includes boxplots visualizing group differences regarding Draft 1 and Draft 2 overall scores. The boxes represent the middle 50% of the data (25th and 75th percentiles). The black squares within the boxes represent group means. The solid horizontal lines through the boxes indicate the median scores. Small black dots mark individual data points (i.e., a single student’s score). The whiskers extending from the boxes show the minimum and maximum scores, excluding the statistical outliers that lay outside the whiskers and are indicated by larger black dots.

Figure 1 shows that both the Dialogic Group and the Traditional Group made progress in their writing performance from Draft 1 to Draft 2. The Dialogic Group has a slightly higher median Draft 2 score ($Mdn = 25.5$) than the Traditional Group, ($Mdn = 24$), a difference found to be statistically significant, $W = 439$, $p = .03^1$. The associated effect size, however, is small ($r = .29$), meaning that although the dialogic feedback group has a statistically significantly higher writing performance, the difference between the groups can be considered small.

Figure 1 also shows that the Dialogic Group has less variation in Draft 2 scores compared to the Traditional Group. For the Dialogic Group, the size of the box representing the middle 50% of the scores became much smaller from Draft 1 to Draft 2. In addition, the whiskers extending from the box, which represent the minimum and maximum scores, became much tighter. That the scores on Draft 2 showed less variation in the Dialogic Group suggests that DF had an equalizer effect in that student scores on Draft 2 were much closer to one another than in Draft 1.

¹ As a reminder, the Mann-Whitney U tests reported in Table 1 showed small ($r = .23$) and non-significant differences ($W = 415.5$, $p = .094$) between the two groups in terms of their overall scores on Draft 1, allowing us to assume equivalence of groups before the treatment and to analyze Draft 2 scores without accounting for Draft 1 scores.

Students' overall writing scores comprised ratings on a range of criteria, including introduction, thesis statement, overview (i.e., content of writing), coherence and cohesion, conclusion, style, grammar, and citations (see Appendix A). Group differences across these criteria were also examined using descriptive and inferential statistics (see Table 2). These differences are also visualized in Figure 2.

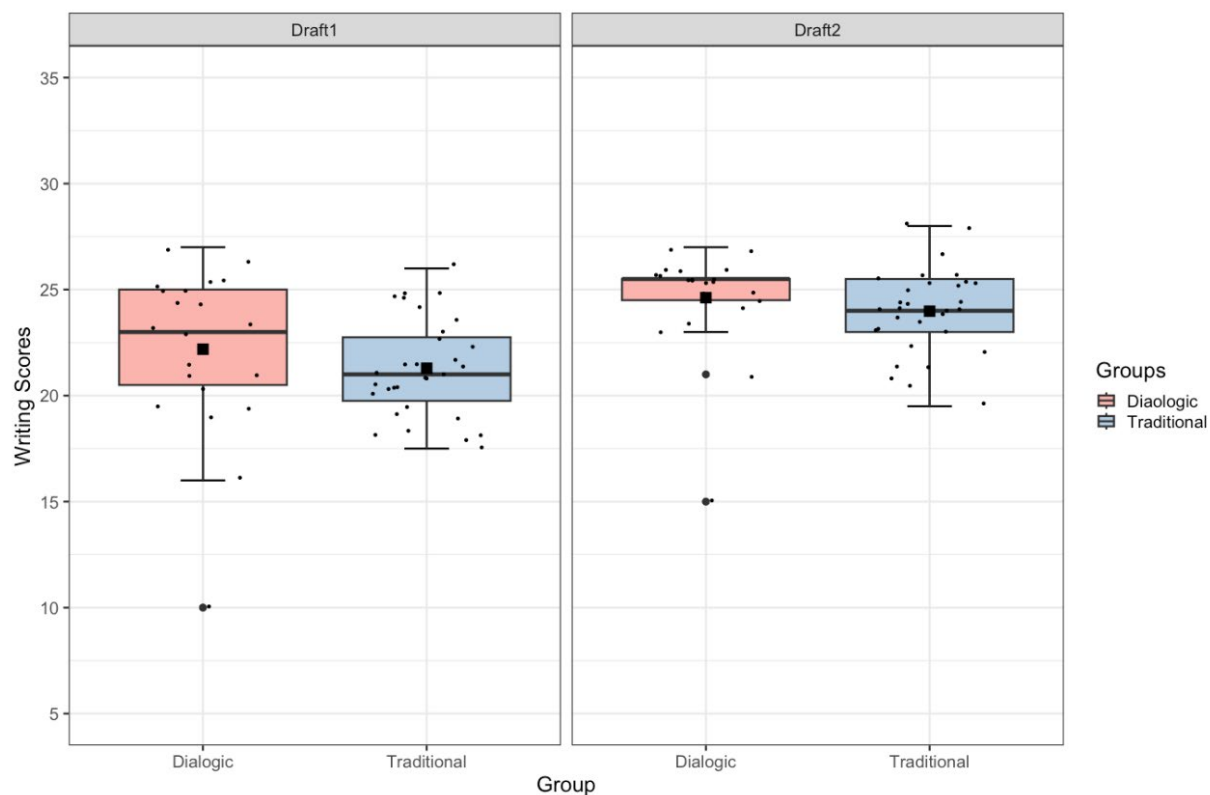


Figure 1 Boxplots Showing Group Comparisons of Total Scores on Draft 1 and Draft 2

Note. The boxplot was generated using the *ggplot2* package (Wickham, 2016) in R (R Core Team, 2026).

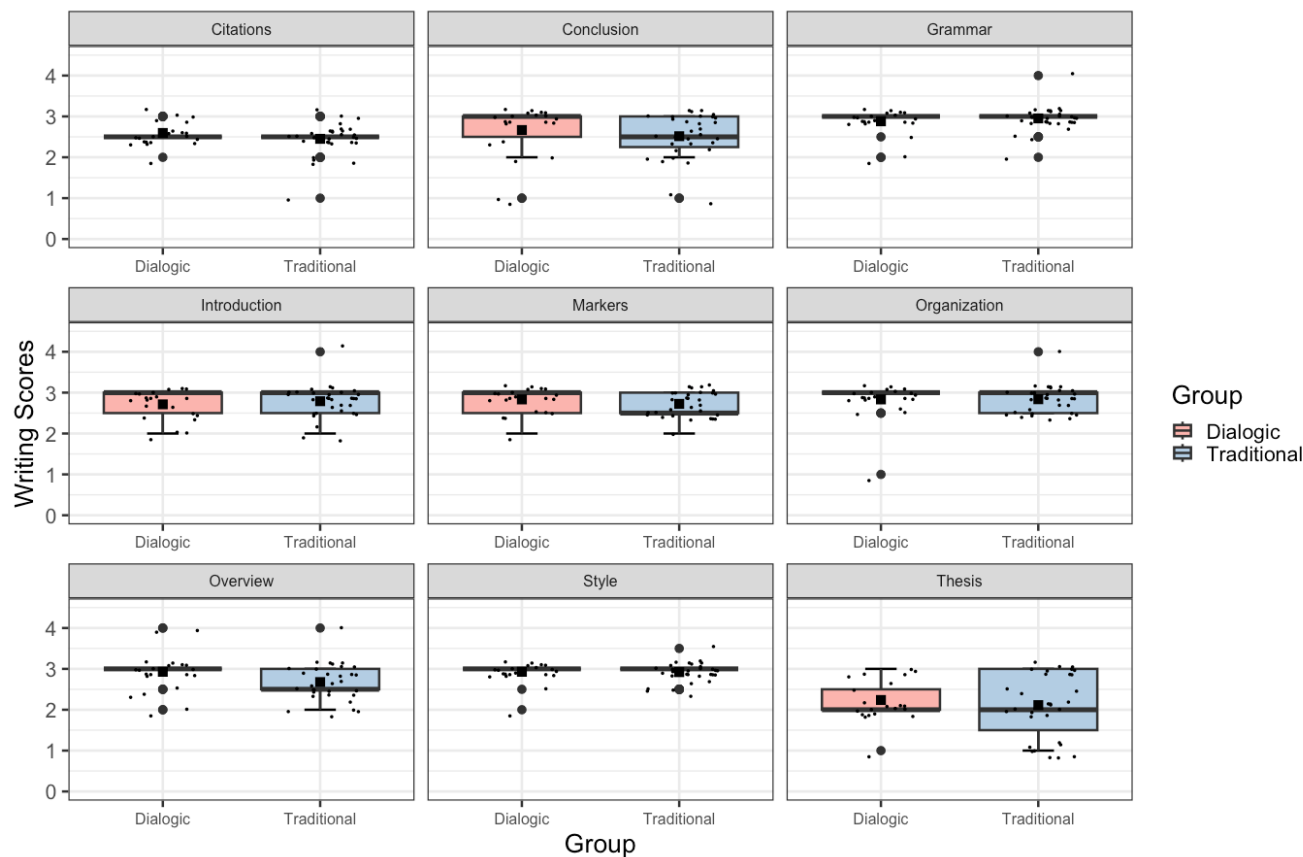


Figure 2 Boxplots Showing Groups Comparisons for Different Aspects of Writing

Table 2 Descriptive and Inferential Statistics Comparing Groups on Different Aspects of Writing

Aspect of writing	Dialogic feedback <i>n</i> = 21		Traditional feedback <i>n</i> = 31		Dialogic vs. Traditional		
	<i>M</i> (<i>SD</i>)	<i>Mdn</i>	<i>M</i> (<i>SD</i>)	<i>Mdn</i>	<i>W</i>	<i>p</i>	<i>r</i>
Overall Score	24.62 (2.58)	25.5	23.98 (1.99)	24	439	*.030	.30 [.03, .55]
Introduction	2.71 (.37)	3	2.79 (.40)	3	301.5	.620	-.07 [-.33, .20]
Thesis	2.24 (.52)	2	2.11 (.78)	2	342	.751	.05 [-.25, .32]
Overview (i.e., Content)	2.93 (.48)	3	2.68 (.44)	3	428	*.037	.29 [.02, .53]
+Organization Gain Score	.29 (.51)	0	.52 (.47)	.5	222	*.039	-.29 [-.54, -.01]
Organization	2.83 (.46)	3	2.84 (.33)	3	375	.265	.16 [-.10, .41]
Markers	2.83 (.29)	3	2.73 (.28)	2.5	395.5	.135	.21 [-.08, .48]
Conclusion	2.67 (.64)	3	2.52 (.56)	2.5	405.5	.103	.23 [-.05, .49]
Grammar	2.88 (.31)	3	2.95 (.30)	3	309.5	.645	-.07 [-.33, .20]
Style	2.93 (.24)	3	2.92 (.23)	3	345	.590	.08 [-.20, .33]
Citations	2.60 (.26)	2.5	2.45 (.37)	2.5	388	.146	.20 [-.07, .44]

⁺Gain scores (Draft 2 minus Draft 1) were used for this particular analysis as group equivalence could not be assumed on the basis of Draft 1 scores.

*Significant at $p < .05$

As can be seen in Table 2, two aspects of writing showed statistically significant differences between the two groups: overview (i.e., content of writing that had to cover the main principles, benefits, and challenges of content-based instruction) and organization. For overview, the Dialogic Group outperformed the Traditional Group with a small effect size, $W = 428$, $p = .037$, $r = .29$. For organization, as previously mentioned, group equivalence could not be assumed on the basis of Draft 1 scores. Thus, gain scores (Draft 2 minus Draft 1) were used in the Mann-Whitney U test, which yielded a statistically significantly larger gain for the Traditional Group with a small effect size, $W = 222$, $p = .039$, $r = -.29$. Despite the greater gain score, it is important to note that after the treatment the two groups had equal median scores ($Mdn = 3$) and nearly equal mean Draft 2 scores for organization (Dialogic Group = 2.83, Traditional Group = 2.84).

4.3 Student Perceptions of Dialogic Feedback

To assess student attitudes towards the implementation of DF and to better inform its future implementation in the course, learner perceptions of mediation were assessed using an open-ended questionnaire. Eighty percent of the comments were positive, 10% were neutral, and 10% were negative. Among these, one response stands out as it raises the issue of the teacher's potential mistrust towards the students: "I reckon we are able to do these steps without having to prove how much we worked on it" (emphasis in the original).

Among those students who commented that the dialogic feedback was a positive learning experience, many (33%) stated that the required DF way of working with revisions was particularly useful. For instance, one student wrote that "It was helpful to track changes and make a list of my changes," while another one explained that "The list of changes made me see what I was missing and what I could do. The tracking changes was for the teacher, so she could see that I changed my paper based on her feedback." In their responses, students expressed their preference for explicit mediations on the part of the instructor as it was believed to enhance their academic writing skills.

The heightened reflection on the writing process and increased awareness of their own strengths and weaknesses was cited by the students as the most important reason why the students perceived the dialogic feedback as useful. One student expressed this sentiment in the following comment: "Having to reflect more on my own paper makes me aware of what I have done is ok and what needs to be changed." The students also pinpointed several times in their comments a better ability to improve writing skills. Students commented that the dialogic process "helps ... learn and improve" and "[allows] to know what was required."

Although most of the students (80%) found the dialogic interaction with the teacher helpful, some (14%) also commented that it was a process that required much work. Nevertheless, the students also admitted that they felt their hard work proved worthwhile, as evidenced by the following comment: "It felt like extra work, but I could see the point of doing it." They also felt it was fair for the teacher to require them to respond to her comments, as illustrated in the following response: "...even though tracking changes and addressing all comments feels like a drag, if you put in the work of improving my text, I should put in some work, too." Thus, while students may have felt that the extra work was time consuming, they recognized the value it had for developing their writing skills.

5. Discussion

The present study adopted a mixed methods design to evaluate the effectiveness of written DF at the tertiary education level and to examine student perceptions of DF in this context. Relative to the first research question, which compared the essay scores of students who participated in a DF task (Dialogic Group) with the essay scores of students who received traditional feedback (Traditional Group), the results indicate that DF can be beneficial for improving overall writing performance, especially the content of writing, but it may not offer advantages for other aspects of writing (e.g., effective thesis statement, organization, style, and grammar). DF can also help reduce variation in student scores as the students in the current study performed more uniformly on Draft 2 when they received DF, as previously illustrated in Figure 1. Taken together, these findings are in line with previous research on DF in that they confirm some positive outcomes of DF (e.g., Alavi & Taghizadeh, 2014; Deng & Sitthitikul, 2025; Jwa, 2025; Maneepakhathorn, 2023; Merkel, 2018; Steen-Utheim & Wittek, 2017) while at the same time cautioning that the extent of potential gains may be limited by factors that were not examined here and thus future research is needed.

Regarding the second research question, most of the students in the Dialogic Group perceived DF as positive and beneficial for the quality of their written texts. Many noted that because they were required to engage in a written dialog about their texts with the instructor, they had to pay close attention to the teacher's comments. This in turn allowed them to become more aware of what they were revising and why, as opposed to if they did not have to respond to the comments. Although some of the students pointed out that the DF was challenging and required substantial work, they admitted that the end result was worth the effort. As the students elaborated, the DF task prompted more in-depth reflection on the writing process and increased their awareness of their own strengths and weaknesses as EAL writers (Jwa, 2025; Maneepakathorn, 2023; Susanti et al., 2025; Turner, 2023). These findings also align with research showing that students do not always understand or make use of written feedback when they engage with it passively (Walker, 2009). Requiring students to respond to comments may therefore help them engage more closely with feedback, even if it also makes the process more demanding.

However, this finding too should be considered cautiously as the findings also revealed some reservations among students about DF as a form of feedback in academic writing. Because some of the student participants expressed their reluctance to engage in a dialogic mediation, and because research on students' perceptions of feedback is scarce, teachers could conduct individual action research projects in their classroom that would help them identify their students' preferred feedback practice. Effective feedback to learners has been acknowledged as a key strategy that contributes to enhancing student learning. The interpretation of *effective* might vary from one student to another and might also not be congruent with their instructor's desired practice(s). Such projects would enable teachers to shed light not only on how feedback should be provided and when but also on how learners might interpret and deal with what they are provided. The concerns that some student participants expressed in this study about DF necessitate future research on student perceptions of effective feedback that would report how feedback is perceived and how it should be utilized by instructors at different grade levels and specific groups. Most significantly, as *dialogic* foregrounds interaction and exchange, we hope that further research will support the inclusion of dialogic feedback practices in teacher-education curricula, promoting assessment approaches grounded in interaction in EAL classrooms.

As an action research project, the current study was limited in its sample size. Small sample sizes reduce the power of statistical tests, possibly leading to Type II errors. That is, the effect of the treatment, namely DF, may not be detected although it may exist. Larger sample sizes could improve this line of research. At the same time, effect size estimation is less susceptible to sample size, which suggests that the effect on other populations of learners could be similar in future studies.

6. Conclusion

This study set out to compare the effectiveness of two forms of feedback on EAL college-level academic writing: DF, in which students engaged in dialogic exchanges with the instructor (Dialogic Group), and traditional feedback consisting of written commentaries (Traditional Group). The Dialogic Group statistically significantly outperformed the Traditional Group in terms of overall writing scores and scores on content of writing, but these group differences were small as indicated by effect sizes. In addition, in comparison to the Traditional Group, the Dialogic Group displayed less variation in scores after the treatment. These findings suggest that DF has the potential to help EAL learners make improvements related to the content of their essays and that it can reduce variance in student performance, allowing more students to succeed. However, the

Traditional Group had larger gains (Draft 2 minus Draft 1 scores) in their organization of the essay, but once again this difference was accompanied by a small effect size, and despite the larger gains, the two groups had almost equal median and almost equal mean Draft 2 scores for essay organization. The study also examined EAL learner perceptions about the DF task. Student responses to an open-ended question suggest that although the participants perceived the task as demanding, they conceded that it benefitted their writing.

Taken together, the findings of this study underscore the importance of moving towards feedback approaches that actively engage learners in the meaning-making process. While the measurable gains associated with dialogic feedback were modest, the combined quantitative and qualitative results highlight its potential to foster deeper engagement, reflection, and learner agency in EAL writing. These outcomes suggest that the value of DF may lie not only in immediate performance improvements but also in supporting longer-term development of self-regulated writing practices. As such, dialogic, interaction-oriented approaches to feedback warrant continued attention in both research and pedagogical practice.

References

- Ajjawi, R., & Boud, D. (2018). Examining the nature and effects of feedback dialogue. *Assessment & Evaluation in Higher Education*, 43(7), 1106-1119. <https://doi.org/10.1080/02602938.2018.1434128>
- Alavi, S. M., & Taghizadeh, M. (2014). Dialogic assessment of writing: The impact of implicit/explicit mediations on L2 learners' internalization of writing skills and strategies. *Educational Assessment*, 19(1), 1-16. <https://doi.org/10.1080/10627197.2014.869446>
- Andrade, H. L., Wang, X., Du, Y., & Akawi, L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *Journal of Educational Research*, 102(4), 287-302. <https://doi.org/10.3200/JOER.102.4.287-302>
- 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>
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17, 102-118.
- Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14, 191-205. <https://doi.org/10.1016/j.jslw.2005.08.001>
- Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400-413. <http://hdl.handle.net/10453/43261>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & evaluation in higher education*, 43(8), 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, 36(4), 395-407. <https://doi.org/10.1080/03075071003642449>
- Carless, D., & Winstone, N. (2023). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*, 28(1), 150-163. <https://doi.org/10.1080/13562517.2020.1782372>
- Chinpakdee, M. (2025). Talking feedback: fostering student feedback literacy through a dialogic approach. *RELC Journal*, 57(1). <https://doi.org/10.1177/00336882251346206>
- Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, L. (2019). What makes for effective feedback: staff and student perspectives. *Assessment and Evaluation in Higher Education*, 44(1), 25-36. <https://doi.org/10.1080/02602938.2018.1467877>

- Deng, Y., & Sitthitikul, P. (2025). The effects of the guided dialogic peer feedback-based writing instruction on Chinese EFL students' writing performance in an integrated blended learning environment. *rEFLECTIONS*, 32(1), 1-27.
- Ene, E., & Kosobucki, V. (2016). Rubrics and corrective feedback in ESL writing: A longitudinal case study of an L2 writer. *Assessing Writing*, 30, 3-20.
<https://doi.org/10.1016/j.asw.2016.06.003>
- Erlam, R., Ellis, R. & Batstone, R. (2013). Oral corrective feedback on L2 writing: Two approaches compared. *System*, 41(2), 257-268. <https://doi.org/10.1016/j.system.2013.03.004>
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83(1), 70-120. <https://doi.org/10.3102/0034654312474350>
- Fernández-Toro, M., & Furnborough, C. (2018). Evaluating alignment of student and tutor perspectives on feedback on language learning assignments. *Distance Education*, 39(4), 548-567. <https://doi.org/10.1080/01587919.2018.1520043>
- Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, 31(2), 315-339. <https://doi.org/10.2307/3588049>
- Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, 32(2), 181-201. <https://doi.org/10.1017/S0272263109990490>
- Ferris, D. R., & Roberts, B. (2006). Error feedback in L2 writing classes: How explicit does it need to be? In P. K. Matsuda, M. Cox, J. Jordan & C. Ortmeier-Hooper (Eds.), *Second-language writing in the composition classroom* (pp. 380-402). Bedford / St. Martins.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Henderikx, M., Vrieling-Teunter, E., Molin, F., & Nadolski, R. (2025). Peer feedback modelling needs in online higher education: An exploratory study. *Studies in Educational Evaluation*, 86, 101464. <https://doi.org/10.1016/j.stueduc.2025.101464>
- Hounsell, D. (2007). Towards more sustainable feedback to students. In D. Boud & N. Falchikov (Eds.), *Assessment in higher education: Learning for the longer term* (pp. 101-113). Routledge.
- Hyland, F., & Hyland, K. (2001). Sugaring the pill: Praise and criticism in written feedback. *Journal of Second Language Writing*, 10(3), 185-212. [https://doi.org/10.1016/S1060-3743\(01\)00038-8](https://doi.org/10.1016/S1060-3743(01)00038-8)
- Hyland, K. & Hyland, F. (2006) (Eds.), *Feedback in second language writing: Contexts and issues*. Cambridge University Press.
- Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144.
<https://doi.org/10.1016/j.edurev.2007.05.002>
- Jwa, S. (2025). Written feedback dialogue: A cyclical model for student engagement with feedback. *TESOL Journal*, 59(2), 1024-1035.
- Li, D., & Qian, J. (2024). Engaging in dialogic peer feedback in L2 writing development: A microgenetic approach. *System*, 126, 103513. <https://doi.org/10.1016/j.system.2024.103513>
- Loewen, S. & Plonsky, L. (2016). *An A-Z of applied linguistics research methods*. Palgrave.
https://doi.org/10.1007/978-1-137-40322-3_1
- Maneepakthorn, R. (2023). An exploratory study of reflective dialogue journal writing of EFL graduate students in the English as a foreign language context. *rEFLECTIONS*, 30(1), 38-58.
<https://doi.org/10.61508/refl.v30i1.264474>
- Merkel, W. (2018). Role reversals: A case study of dialogic interactions and feedback on L2 writing. *Journal of Second Language Writing*, 39, 16-28.

- Molloy, E., & Boud, D. (2013). Changing conceptions of feedback. In D. Boud & E. Molloy (Eds.), *Feedback in higher and professional education: Understanding it and doing it well* (pp. 11-33). Routledge.
- Muncie, J. (2000). Using written teacher feedback in EFL composition classes. *ELT Journal*, 54(1), 47-53. <https://doi.org/10.1093/elt/54.1.47>
- Nurmukhamedov, U. & Kim, S.H. (2010). ‘Would you perhaps consider...’: hedged comments in ESL writing. *ELT Journal*, 64(3), 272-282. <https://doi.org/10.1093/elt/ccp063>
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129-144. <https://doi.org/10.1016/j.edurev.2013.01.002>
- Perpignan, H. (2003). Exploring the written feedback dialogue: A research, learning and teaching practice. *Language Teaching Research*, 7(2), 259-278.
- Plonsky, L., & Oswald, F. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878-912. <https://doi.org/10.1111/lang.12079>
- R Core Team (2026). R: A language and environment for statistical computing [Computer software]. <https://www.R-project.org/>
- Steen-Utheim, A., & Wittek, A. L. (2017). Dialogic feedback and potentialities for student learning. *Learning, Culture and Social Interaction*, 15, 18-30.
- Susanti, A., Sharif, T. I. S. T., & Masardo, A. (2025). Dialogic feedback in English Language Teaching (ELT): Modalities, outcomes, and research trends. *English Language Teaching Educational Journal*, 8(1), 37-53.
- Sugita, Y. (2006). The impact of teachers’ comment types on students’ revision. *ELT Journal*, 60(1), 34-41. <https://doi.org/10.1093/elt/cci079>
- Tam, A. C. F. (2021). Undergraduate students’ perceptions of and responses to exemplar-based dialogic feedback. *Assessment & Evaluation in Higher Education*, 46(2), 269-285. <https://doi.org/10.1080/02602938.2020.1772957>
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.
- Truscott, J. (1999). The case for “the case for grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing*, 8(2), 111-122.
- Truscott J. (2007) The effect of error correction on learners’ ability to write accurately. *Journal of Second Language Writing*, 16(4), 255-72.
- Turner, E. (2023). Dialogic feedback and literary disciplinary knowledge in L2 writing instruction: how attitude to feedback influences academic achievement. *Research Papers in Education*, 38(1), 21-44. <https://doi.org/10.1080/02671522.2021.1941216>
- Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.
- Walker, M. (2009). An investigation into written comments on assignments: do students find them usable?. *Assessment & Evaluation in Higher Education*, 34(1), 67-78. <https://doi.org/10.1080/02602930801895752>
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.
- Wickham, H. (2016). *Ggplot2: Elegant Graphics for Data Analysis*. Springer.
- Winstone, N. E., & Carless, D. (2019). *Designing effective feedback processes in higher education. A learning-focused approach*. Routledge.
- Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners’ agentic engagement with feedback: A systematic review and taxonomy of recipience processes. *Educational Psychologist*, 52(1), 17-37. <https://doi.org/10.1080/00461520.2016.1207538>
- Yang, M., & Carless, D. (2013). The feedback triangle and the enhancement of dialogic feedback processes. *Teaching in Higher Education*, 18(3), 285-297. <https://doi.org/10.1080/13562517.2012.719154>

Appendices

Appendix A

Assessment Rubric

Criteria	Outstanding 4	Good 3	Adequate 2	Weak 1	Unacceptable 0
Introduction: Strong introduction with appropriate context that raises the overall topic and sets the stage for the remaining paper in an engaging way					
Thesis: Clear thesis statement that reflects the goal of the essay					
Overview: Clear and concise overview the main principles, benefits, and challenges of CBI					
Organization: Effective organization (logical division into paragraphs). The essay has an introduction, body and conclusion.					
Markers: Explicit markers of coherence and cohesion (appropriate conjunctions, transition sentences, topic sentences)					
Conclusion:					

Effective conclusion that does more than simply sum up paper					
Grammar: The essay is written in grammatically correct, academic English, including correct punctuation and capitalization.					
Style: The writer uses a clear, concise style with a variety of lengths and types of sentences.					
Citations: In-text citations provided in APA style					
TOTAL SCORE					

Appendix B

Shapiro-Wilk Normality Tests

	Dialogic feedback <i>n</i> = 21		Traditional assessment <i>n</i> = 31	
	<i>W</i>	<i>p</i>	<i>W</i>	<i>p</i>
Introduction 1	.80885	.0008988	.83924	.0003014
Introduction 2	.72708	6.107e-05	.78562	2.828e-05
Thesis 1	.76447	.0001971	.78666	2.952e-05
Thesis 2	.76447	.0001971	.82495	.000155
Overview 1	.81513	.001128	.81978	.0001226
Overview 2	.77041	.0002395	.83735	.0002756
Organization 1	0.72686	6.067e-05	0.80578	6.616e-05
Organization 2	.42178	4.171e-08	.6853	7.104e-07
Markers 1	.77708	.0002988	.79217	3.71e-05
Markers 2	.6181	3.077e-06	.71704	2.094e-06
Conclusion 1	.82309	.001512	.80735	7.082e-05
Conclusion 2	.58934	1.522e-06	.78914	3.269e-05
Grammar 1	.62932	4.084e-06	.76971	1.49e-05

Grammar 2	.42779	4.686e-08	.54382	1.127e-08
Style 1	.69107	2.135e-05	.76971	1.49e-05
Style 2	.34311	9.688e-09	.61005	6.961e-08
Citations 1	.82021	.001359	.86386	.001014
Citations 2	.68264	1.686e-05	.6784	5.665e-07
Overall score 1	.87159	.01024	.95961	.2849
Overall score 2	.66009	9.115e-06	.97257	.5922

Notes. 1 = Draft 1; 2 = Draft 2. Normal distribution is assumed when $p > .05$. Cells with $p > .05$ are shaded gray. There are no comparisons for which both group data are normally distributed; thus, Mann-Whitney U tests were used instead of independent samples t-tests.