

Scaffolded Peer Feedback and Second Language Writing Proficiency: Implications for Inclusive ESL Teaching Practice

February 2023 – Volume 26, Number 4

<https://doi.org/10.55593/ej.26104a8>

Carlo Domingo Casinto

Bahrain Polytechnic, Bahrain

<carlo.casinto@polytechnic.bh>

Abstract

Although studies on peer feedback have identified several advantages for second language writing classes, the benefits of peer feedback to low-proficiency students have not been thoroughly investigated. This current study aims to determine how low-proficiency second-language students of English can improve their writing by giving and receiving peer feedback. Utilizing mixed methods, data collection and analyses were primarily conducted through posttest-only control-group design and focus group discussion (FGD). There were twenty university student participants randomly selected from two writing mixed-proficiency level classes. Framed under Vygotsky's sociocultural theory, an analysis of the gains in writing ability measured from formative and summative writing samples indicated that the low-proficiency students, specifically in the experimental group, made more significant gains in their writing over the course of the semester than the high-proficiency students in both groups. Even though the results generally showed no significant difference, except for higher error avoidance efficiency between experimental and control groups and among the high-proficiency students in both groups, a significant difference was observed between the low-proficiency students in both groups (control group $M = 81.6$, experimental group $M = 89.8$, control group $SD = 8.3$, experimental group $SD = 9.5$, $t(4) = -3.23$, $p = .032$). This leads to the conclusion that low-proficiency students benefit from scaffolded peer feedback.

Keywords: Scaffolded peer feedback (SPF); Peer review; Peer editing; Second language writing

Writing, as a productive language skill, is one of the most challenging abilities to teach in a second language (L2) class. Peer review, as a teaching approach in writing, enables teachers to give their students more feedback in terms of vocabulary and word expressions, grammar, structure, and content. These are the local and global aspects of writing that most L2 learners struggle with and need more support on. Moreover, peer review affords L2 learners an

opportunity for meaningful interaction with other learners, active collaboration with peers such as brainstorming on content and ideas, and critical discussion on the writing itself such as editing word usage, cohesion, and organization.

However, some challenges confront peer review in the writing process. First, there is an attitude of distrust (Alsehibany, 2021). Students perceived peer review as a counterproductive classroom activity, as some peer reviewers are seen as less competent than the writers (Saeli & Cheng, 2021). As reported by Salih (2013), generally peer review tends to be focused on grammar rather than content, ideas, structure, organization, and cohesion. Teachers, on the other hand, perceived it as having no impact on writing quality because most students do not stay on task during the activity (Brammer & Calera, 2007). Because of these challenges, the pedagogical value of peer review in the L2 classroom is questioned by some critics.

Nevertheless, the idea that peer review can be a useful teaching approach in L2 writing is advanced by many studies (Bagheri & Rassaei, 2022; Kim, 2012; Lundstrom & Baker, 2009; Min & Chiu, 2021; Saeli & Cheng, 2021). This is particularly true when students are given sufficient training on how to give scaffolded peer feedback (SPF) (Alnasser, 2018; Alsehibany, 2021; Do, 2020; Leijen, 2017; Sippel, 2021; Zhao, 2018). SPF can be integrated into the writing process promoting a student-centered L2 classroom where learners are empowered to critically evaluate their writing (Brammer & Calera, 2007; Nassaji, 2016; Tajabadi et. al, 2020). SPF sessions can provide L2 learners with a social platform to interact with an audience. Through SPF, L2 learners can see the perspectives of others apart from their own (Zhu, 2001), and negotiate and collaborate on improving their writing (Nguyen, 2021; Tajabadi, et. al, 2020).

Finally, SPF can stimulate L2 students, especially those with low proficiency and those who are not confident with their language abilities, to develop a more positive attitude toward writing as SPF can serve as their classroom social support system in accomplishing the language task (Bolourchi & Soleimani, 2021).

Literature Review

The potential benefits of peer review to low-proficiency L2 students are the dimensions that have rarely been studied empirically in L2 writing research. For instance, the ability to critically evaluate writing and provide effective feedback, especially on a global level such as content and organization as essential writing skills is understudied (Sotouddehnama & Pilehvari, 2016).

One aspect of peer review that should be investigated is how it improves the evaluation skill of L2 learners. Evaluation skill enables L2 students to develop their critical eyes and see logical gaps and problems with cohesion and organization. They are also able to identify other local issues that impact the clarity of their own and their classmates' writings (Lundstrom & Baker, 2009; Zhao, 2018). In the process, as reported by Hu and Lam (2010), the dyadic and collaborative interactions involving learners regardless of their levels of proficiency during the peer review would lead to developing stronger writing competency and better writing quality because they exchanged mutually beneficial feedback.

This occurrence in second language learning is explicated by Vygotsky's sociocultural theory which postulates that learners' ability to obtain and process information is limited to their zone of proximal development (ZPD) or the space between what a learner can do independently and what a learner can do with guidance and assistance (Lantolf & Beckett, 2009). However, with

peer review, the learner's ZPD is expanded, and they can perform a task on their own interdependently with other learners (Hossein & Eva, 2017). Thus, this current study is anchored on Vygotsky's sociocultural framework which posited that language acquisition is facilitated through learners' social interaction in which they negotiate and elucidate to confirm, clarify, or recast, and comprehend information mutually given and received (Vygotsky, 1986).

This communicative interaction between learners can provide them with scaffolded support to give feedback especially when it is focused on a particular language form during the negotiation process (Nassaji, 2016). Consequently, this expands learners' zone of proximal development (ZPD) specifically those with low-proficiency and enables them to improve comprehension and acquire the target language subsequently (Vygotsky, 1986). When ZPD is expanded, learners can also improve their uptake or their ability to respond after the feedback (Nassaji, 2016). This response is a learner's action to repair or correct the errors in the original output following feedback. Hence, social interaction is the baseline from which learners' ZPD can be enlarged to clear the way for language acquisition.

Benefits and Limitations of SPF

Previous studies elucidate that the benefits of peer review are derived from joint scaffolding or the extra support that learners provide to help each other. For instance, Zheng (2012) demonstrates how two students with different levels of L2 proficiency learn from each other during peer review sessions. In the analysis, the students demonstrate how their dual roles in the peer review process, such as serving as a reviewer and as a writer, enable them to scaffold the learning of their peers. Hence, as asserted by Umaz (2019) and Tajabadi et. al, (2020), it can be contended that even low-proficiency students can scaffold their peers' learning, and when they can do that, it may benefit them in their writing.

Consequently, Sotouddehnama and Pilehvari (2016) and Kim and Emeliyanova (202) explicated that an L2 student who can critically evaluate others' writing may develop self-review proficiency, or a learner's ability to probe into their writing and objectively identify areas that need improvement. This is so because scaffolding requires mainly a collaborative pattern of interaction between the reviewer and the writer (Storch & Suzuki, 2020). The type of feedback generated from a collaborative interaction is the outcome of the negotiation, recast or direct correction, clarification requests, and elicitation prompts (Nassaji, 2016). Zhao (2018) expounded that the negotiation process of a scaffolded peer review facilitates the development of a set of mediating strategies that includes clarification, justification, and confirmation of the feedback given.

Based on the foregoing studies, L2 learners can potentially develop their sociolinguistic competence in the peer review process. When L2 students learn to assess and evaluate their own and their peers' papers and to identify areas that require improvement, Fatoni and Tauchid (2020) and Nguyen (2021) elucidated that it is a categorical indicator of the positive value that peer review brings to the whole writing process.

Although there has been sufficient empirical evidence pointing towards the positive impact of peer review among low-proficient L2 learners, the study of Kamimura (2006) revealed that peer review activities are more advantageous to highly efficient L2 learners who exhibit greater improvements in the overall quality of their writings. This study, nevertheless, does not contradict the preceding findings as it does not refute the benefits of peer review to low-proficiency students.

In addition, some studies cited the limited sociolinguistic value of peer review. In the studies of Salih (2013) and Saeli and Cheng (2021), critics argued that low-proficiency students tend to focus more on the micro/local aspects of writing (e.g., grammar, vocabulary, word choice) instead of its macro/global aspects (e.g., content, organization, cohesion), which would prevent them from developing essential critical evaluation skills. Conversely, the studies of Sanchez-Naranjo (2019), Do (2020), and Bolourchi and Soleimani (2021) explained that these issues confronting the pedagogical value of peer review can be addressed by proper training and sufficient scaffolding.

There are several requisites in developing the ability to scaffold feedback that must be addressed by training. According to Alnasser (2018), Sippel (2021) and En-Chong (2022), these include providing form-focused instructions such as evaluation forms and checklists, and setting up a simulation or a mock review that models the process by which sociolinguistic strategies are employed consistently throughout. Unless these conditions are satisfied, L2 students may not be able to scaffold through their peer feedback since the collaborative pattern of interaction is not established to stimulate negotiation and mediation (Brammer & Calera, 2007; Chang, 2016).

Training as a pre-conditioned to an effective SPF

Training, therefore, is indispensable in SPF (Leijen, 2017). In the studies of Zhao (2018) and Tajabadi, *et. al.*, (2020) questions emerged, however, as to whether training can foster a collaborative pattern of interaction between peers. In addition, the study of Sato and Lyster (2012) on Japanese EFL learners suggests that training can effectively nurture this kind of interaction during the review process. Remarkable improvements in accuracy and fluency in terms of the ability to recast and provide prompts are observed. In two separate studies of peer review training on L2 learners in Saudi Arabia, an in-class modeling of good peer review practices was implemented in which an instructor provided students a "walk-through" or a class demonstration on how to evaluate a draft using a prepared evaluation form or checklist. Then, a simulation exercise or practice was conducted. It was found that peer feedback had improved its quality and coverage both at the micro/local level (e.g., grammar, mechanics) and at the macro/global level (e.g., idea, content, organization) (Alnasser, 2018; Alsehibany, 2021). Moreover, in two separate experimental studies of peer review training on L2 learners in Vietnam and in New Zealand, it was observed that learners achieved a higher total gain score and developed positive attitudes task completion, and ideas development respectively (Do, 2020; Hislop & Stracke, 2017). Hence, this current study intends to incorporate training in its investigation of SPF on low-proficiency students.

Research Problems

This current study investigates whether low-proficiency L2 learners who have been given social support through the scaffolded review process would develop sociolinguistic competence and higher proficiency in L2 writing. Specifically, it will address the following questions:

(RQ1) Which group and proficiency level of L2 students have a higher average gain score and a higher error avoidance efficiency rate in their writing?

(RQ2) If low-proficiency students in the experimental group do improve their writing ability more than those in the control group, which writing aspects (both global and local) do they improve?

(RQ3) Is there a significant difference in the writing mean scores of control and experimental groups and between low-proficiency students and high-proficiency students in both groups?

(RQ4) What are the recurring themes in SPF as perceived and experienced by the students in the experimental group?

(RQ5) What are the implications of these questions to ESL teaching?

Methodology

Participants

Twenty students enrolled in two classes of English Communication for Research, Reading and Writing in the Discipline course offered in semester two of the academic year 2020-2021 of a co-educational university in the Kingdom of Bahrain participated in this study. The number of participants represented a 50% sampling rate of the cohort population. They were chosen based on either of the following criteria: having a CEFR (Common European Framework of Reference for Languages) level at B1 or lower or must have a level at B2 or higher. The other essential criteria for participants in the experimental group were: having attended the SPF training sessions (demonstration and simulation); having reviewed three peer's formative papers; having received feedback from three formative papers; having attended the post-review conference; and having done three paper revisions.

For this study, B1 or lower level was categorized as a low-proficiency L2 learner while B2 or higher level was categorized as a high-proficiency L2 learner. The language proficiency levels of the participants of this study were generated from their assessment records at the School of Foundation of the university. There were ten participants of mixed levels in each class. To eliminate proficiency bias in each class, five participants had B1 level while the other five had B2 level. After the equal number of mixed-level participants was ascertained, the experimental and the control groups were randomly determined through a lottery of the two EL5002 classes.

Research and Ethical Procedures

Under the Posttest-Only Control-Group Design, the participants who expressed in writing their prior, free and informed consent to participate in the study were randomly assigned and divided into the experimental group ($n=10$) and the control group ($n=10$). All participants in both groups were given alphanumeric codes as their unique identifiers to protect their personal information and to ensure their confidentiality and privacy.

	Group	Pre-test	Intervention	Post-test	Outcome
Random Assignment of Participants	Experimental	None	Scaffolded Peer Review (SPF)	Yes	L2 Writing Proficiency
	Control	None	None	Yes	

Figure 1. Posttest-Only Control-Group Design Flow Chart

Following strictly COVID-19 safety protocols for face-to-face learning modality, on the second week of the semester, participants in the experimental group were given a series of SPF training sessions. The first session was an orientation and a demonstration on how to use the checklist and the evaluation form that covered global and local aspects of writing. The second session was a simulation or mock exercise in which the participants were immersed in the SPF process.

In the fourth week, participants from both groups started to work on the first of the three formative writing drafts after receiving similar lessons on writing in the preceding weeks. The lessons included how to write the background and the thesis statement in the introduction, comparison and contrast analysis in the body paragraphs, and the CERA (claim, evidence, reason, and action) framework in the conclusion. This is to ensure that all participants in both groups receive the same amount and level of instruction, except for the SPF procedure which only the experimental group had done. Thus, the formative writing activities required that all participants had to listen and take down relevant notes from an audio text at the target course exit C1 level. Based on their notes, they were tasked to write a 300 to 350- word four-paragraph expository essay utilizing mainly persuasive writing methods, namely: comparison and contrast analysis and CERA within a 90-minute time limit.

This was followed by a peer review using the SPF process during the second session for those participants in the experimental group. The participants were purposively paired according to their proficiency level, which requires pairing of B1 level and B2 level participants (Storch & Aldosari, 2012; Zheng, 2012). The first hour of the two-hour SPF session was focused on giving feedback on the global or macro aspects and local or micro aspects of the writing drafts. The subsequent hour was for the peers to meet in a post-review conference for them to negotiate and mediate meanings, to clarify, justify, and confirm content and language, and to elicit and prompt ideas from each other before they act on the feedback for revisions (Nassaji, 2016; Zhao, 2019).

On the other hand, during the second session of the fourth week, the participants in the control group reviewed and revised their papers on their own guided only by the checklist. By depriving the participants in the control group of SPF sessions, this current study was able to investigate the research questions of whether SPF benefits low-proficiency L2 students' writing.

Writing on two other different topics, the same procedures were also conducted in two separate formative writing sessions conducted on weeks eight and twelve respectively. Subsequently, on the fifteenth week, the summative assessment which served also as the posttest to measure writing proficiency the dependent variable for this study was conducted for both groups.

To guarantee the integrity and reliability of the test results, all submissions by both groups were scanned through a plagiarism checker utilizing software installed as a plugin in the learning management system of the course. Submissions with a similarity rating of 26%-35% are penalized with a 5% deduction from the task fulfillment band score. Moreover, a similarity rating of 36% or higher is outrightly rejected. No submission from the participants was meted with any of the penalties.

In the sixteenth week, a series of focus group discussions (FGD) of the participants in the experimental group was held. To eliminate hesitation among low-proficiency participants and domination of opinion by strong participants, two separate sessions for the group based on their

proficiency levels were conducted. Each session lasted for an hour eliciting participants' experiential opinions based on the perceived strengths and weaknesses of SPF.

The FGD sessions observed the following guidelines: First, the facilitator introduced the FGD and its rationale and objectives; second, the participants were asked to introduce themselves by stating their names and their ID numbers; third, the participants were asked about their SPF experience: during the peer review and the peer review conference. Each participant was given time to speak and the opportunity to agree or disagree with the other participants' opinions. On the other hand, the facilitator took cues from the explanation of each participant and drew more insights and probing responses. Finally, the facilitator summed up the proceedings and asked if participants had other things to add. After this, the facilitator thanked the participants and closed the FGD.

Instruments, Assessors, and Markings

To ensure the validity of the summative assessment or posttest, it had to conform with the learning outcomes, task requirements, time limit, and the exit level of the course. It was similar to the three writing drafts that participants in both groups had written during the formative sessions. As stated, the assessment required that participants had to listen and take down relevant notes from an audio text at the target exit C1 level. Based on their notes, they were tasked to write a 300 to 350- word four-paragraph expository essay utilizing mainly persuasive writing methods, namely: comparison and contrast analysis and CERA within a 90-minute time limit (Please see Appendix A).

The posttest was developed and moderated by tenured English language faculty members of the university in semester two of the academic year 2020-2021. Two of the faculty members, who were not involved in moderation, served as assessors. Before marking, the assessors conducted a norming session to standardize the interpretation of the writing rubric. Double-blind average marking was conducted for the twenty papers of participants from both groups. A third assessor from the faculty was required in instances when there was a discrepancy of more than five marks between the first and the second assessors. The average of the three ratings was taken as the writing grade in this case.

To assess the writing proficiency of the students, the writing rubric was developed by the faculty of English language of the university (Please see Appendix B). There were nine criteria in the rubric covering the learning outcomes at the CEFR C1 target exit level. The rubric was used to assess both the local and the global aspects of formative and summative writing assessments. It was based on a five-point scale per criterion, and student essays were assigned a score on the scale on the following writing aspects: task fulfillment (criterion 1); organization, cohesion, and coherence (criterion 2); grammar (criterion 3); CERA structure (criterion 4), vocabulary (criterion 5), tone and register (criterion 6), punctuation and capitalization (criterion 7); spelling (criterion 8); and word count (criterion 9).

At the highest point of the scale, task fulfillment refers to the content coverage that has at least six (6) accurate and relevant points from the listening audio text (three for each body paragraph: one similarity and two differences) and effectively paraphrases this information with similarity rating of 20% or below. Secondly, organization, cohesion, and coherence refer to an effective introduction, the body paragraphs, and recommended course of action; the relationships between ideas and content effectively indicated; a wide range of cohesive devices used effectively; and the variety of referencing. Thirdly, grammar refers to the accurate use of a

broad range of appropriate grammatical structures (e.g., passive voice and complex sentences). Fourthly, the CERA structure refers to the relevant claim given and supported with adequate evidence and reasoning in the recommended next course of action paragraph. Fifth, vocabulary refers to a wide range of word usage that is appropriate and accurate. Sixth, tone, and register refer to appropriateness to the audience and purpose consistently, no inappropriate use of slang or idiomatic expressions, and no contractions. Seventh, punctuation and capitalization refer to no mechanical errors in writing. Eighth, spelling refers to no spelling error in writing. Finally, word count refers to the number of words in writing by the set limit. The criteria scores on the rubric were then added and converted into percentage score which is the overall writing grade.

Conversely, the error avoidance efficiency rating (EAER), which in theory refers to the learners' uptake of each participant in both groups, is determined by the local and global errors committed during the series of formative assessments that were coded together and compared with those in the summative assessment. The difference between recurring and corrected errors expressed in percentage scores refers to the error avoidance efficiency rate of each participant. Moreover, the groups (experimental and control) and the clusters (B1 and B2) EAER averages were also ascertained to serve one of the research objectives of this current study (Please see Appendices G and H).

Furthermore, to warrant the validity of the rubric, it underwent pre-moderation and post-moderation. The first was to have senior tenured language faculty members examine the constructive alignment of each criterion in the rubric to the learning outcomes, the assessment task, performance, and the target exit level of the course. The second was to have two of the faculty members, who were teaching the course, pilot test the rubric in the first session of the formative assessment. Ambiguous scaled-criteria descriptors in the rubric were noted, corrected, amended, or revised collectively in a post-moderation meeting. Subsequently, before marking, the assessors conducted a norming session to regulate the interpretation of the writing rubric. In the session, sample essays that were deemed high, average, and low were randomly chosen and rated by each assessor. Then, the assessors compared their scores and discussed the reasons for their ratings until all of them agreed on a given score. This process was repeated for all other criteria in which there were disagreements on their rating scores.

To scaffold the peer feedback process, a review checklist and an evaluation form were developed by the proponent of this current study. Both SPF documents were designed according to the assessment criteria in the rubric, and pilot-tested in two phases (Please see Appendix C). The first phase of the pilot test was by the tenured faculty members of the English Program of the university. The second phase of the pilot test was by the participants of the experimental group during the simulation session on the second week. Equivocal areas in both documents were identified and revised. Consequently, the final drafts of the two SPF documents were rolled out on the fourth week which was the first of the three formative session-week incorporating SPF as part of the writing process. The checklist prompted participants in the experimental group to examine their peer's writing and to look for evidence in terms of content, organization, and cohesion requirements to support their feedback. On the other hand, the evaluation form elicited from the participants their qualitative opinions about the quality of writing of their peers and how their peers can improve on their writing. Their feedback was supported by evidence indicated in the checklist. The two forms were the basis of the post-review conference.

Finally, to facilitate the focus group discussions (FGD) of the participants of the experimental group, an FGD guide was developed by the proponent of this study and validated by the manager of the English program of the university. The guide contained open-ended questions inquiring about the participants' perception of their experience of the SPF process, namely: during the review and during the conference (Please see Appendix D). During the review, the participants were asked about how reviewing their peers' writings had helped them review and revise their writing. On the other hand, during the conference, the participants were questioned about how discussing their feedback with their peers stimulated them to reflect on their writings but from a different perspective.

Results and Discussion

The first question of this study inquired on which group and proficiency level of L2 students had a higher average gain score and a higher error avoidance efficiency rate in their writing. To answer the sub-question on the gain score, the average scores in the three formative assessments were compared with the summative scores of both groups which were clustered according to their proficiency levels. Moreover, to answer the sub-question on error avoidance rate, the coded errors identified in the three formative assessments were compared to the coded errors in the summative assessments.

These comparative analyses served two purposes: First to determine how much gain scores were achieved by both groups and their respective clusters, and second to ascertain how efficiently the errors were avoided by both groups and the clusters. The results of these analyses revealed that the mean gain score of the experimental group at 10.91% is higher compared to the mean gain score of the control group at 6.90%. Specifically, the mean gain score of low-proficiency students in the experimental group at 13.81% was higher than the mean gain score of low-proficiency students in the control group which registered at 9.20%. In addition, the mean gain score of high-proficiency students in the experimental group at 6.30% was higher than the mean gain score of high-proficiency students at 4.61% (Please see Appendices E and F).

In the aspect of error avoidance efficiency rate (EAER), the experimental group had 23.78% while the control group had a 13.33% efficiency rating. In terms of proficiency level, the EAER of low-proficiency students in the experimental group at 22.57% was higher than the EAER of low-proficiency students in the control group at 6.66%. Similarly, the EAER of high-proficiency students in the experimental group at 25% was higher than the EAER of high-proficiency students in the control group at 20%. These findings supported the contentions of Do (2020) and Alsehibany (2021) who both asserted that SPF had helped students achieve higher gain scores, more efficient task completion, and better writing quality (Please see Appendices G and H).

The second question of this research study probed whether low-proficiency students in the experimental group did improve their writing ability more than those in the control group and on which writing aspects (both global and local) did they improve. An inventory of the errors avoided revealed that the low-proficiency students were able to improve relatively on content, spelling, vocabulary, lexical semantics, and syntax. Conversely, recurring errors were noted in organization, cohesion, and complex and run-on sentences. By contrast, the low-proficiency students in the control group had a dismal improvement which was limited to spelling and vocabulary or word choice. This explained the wide gap at 15.84% in the EAER of B1 clusters

in both groups. These findings confirmed the assertions of Lundstrom & Baker (2009), Fatoni & Tauchid (2020), and Saeli & Cheng (2021) who contended that SPF had supported students, especially those at the lower proficiency level, made significant gains in most aspects of their writing.

Table 1. Significance of the Difference in Writing Proficiency according to Group

Group	<i>n</i>	<i>M</i>	<i>SD</i>	<i>cv</i>
Control	10	85.7	7.2	-2.26
Experimental	10	89.3	7.1	

The third question queried whether there was a significant difference in the writing mean scores of control and experimental groups and between low-proficiency students and high-proficiency students in both groups. To examine whether there was a significant difference between groups and between clusters in each group, a *t*-test for paired two samples for means was run for group and each proficiency cluster in each group. The group test yielded $t(9) = -1.63, p = .137$ which is greater than the .05 level of significance for this study. This meant that there was no significant difference in the writing proficiency of L2 students in groups. This finding generally contradicted most of the literature cited in this current study, specifically from Uymaz (2019) and Kim and Emeliyanova (2021) who both reported a positive correlation between SPF and writing proficiency. Nonetheless, it supported the contention of Bagheri and Rassaei (2022) who asserted that although students had benefitted from SPF, the difference was not statistically significant.

However, it must be mentioned that this current study was delimited to the following factors: First, the pairing was heterogenous (high-low proficiency level) not homogenous (high-high and low-low proficiency levels). Second, the sample size was relatively small and limited only to one CEFR proficiency level (B1 and B2). Third, the formative assessment was limited to three drafts and covered only one genre (expository essay). Finally, the participants covered only L2 students. Either of these factors may have influenced the statistical outcome of this study.

Table 2. Significance of the Difference in Writing Proficiency according to B1 Cluster

Cluster	<i>n</i>	<i>M</i>	<i>SD</i>	<i>cv</i>
B1 Control	5	81.6	8.3	-2.78
B1 Experimental	5	89.8	9.5	

The test for B1 cluster demonstrated $t(4) = -3.23, p = .032$ which was less than the .05 level of significance for this study. This meant that there was a significant difference in the writing proficiency of L2 students within the low proficiency level. This suggested that the B1 cluster in both groups had improved in some aspects of the writing process, but the B1 experimental cluster had greater gains than the other cluster on some of these aspects. Several plausible reasons can be inferred from this finding.

One possible reason was that since the language skills of low-proficiency students were less developed, they had more room for sociolinguistic improvement because of the peer interaction than did the high-proficiency students; hence, those skills they developed enabled them to achieve greater relative improvement in their writing ability. Nevertheless, this finding supports the proclamations of Lundstrom and Baker (2009), Fatoni and Tauchid (2020), and Saeli and Cheng (2021) who declared that SPF had supported low-proficiency students made significant gains in both global and local aspects of their writing, most notably in terms of their lexical semantic and syntactic skills. In contrast, this finding confirmed the contradiction in the argument of Kamimura (2006) who claimed that SPF had benefitted more high-proficiency students than low-proficiency students.

Table 3. Significance of the Difference in Writing Proficiency according to B2 Cluster

Cluster	<i>n</i>	<i>M</i>	<i>SD</i>	<i>cv</i>
B2 Control	5	89.8	2.1	2.78
B2 Experimental	5	88.9	4.8	

The test for B2 cluster demonstrated $t(4) = 0.361, p = .736$ which was less than the .05 level of significance for this study. This meant that there was no significant difference in the writing proficiency of L2 students within the high proficiency level. This finding contradicted again Kamimura (2006) who asserted that high-proficiency students benefitted more from SPF than low-proficiency students. The assertion was untenable in this current study.

This finding could be inferred to mean that since the language skills of high-proficiency students were relatively more advanced, the improvement they made through the SPF process seemed less significant relative to where they began. Furthermore, high-proficiency students may be developing skills that would take longer than one semester to develop; thus, the benefits would not manifest during this current study. However, this finding supported the contention of Lundstrom and Baker (2009) who viewed the benefits of SPF mainly on the qualitative aspects of writing as evidenced by the higher mean gain scores and the higher EAER that B2 level students in the experimental group had nonetheless achieved.

Generally, the findings indicated that although students in both proficiency levels in the experimental group had higher gain scores and higher EAER, they did not have significant differences in their writing proficiency, except those in B1, the students in the low proficiency level. Nevertheless, notable improvements were observed in the areas of content, organization, cohesion, word usage, and syntax in the writing of students in the experimental group, particularly those at the B1 level. Thus, while further research was needed to validate these findings, it appeared that the observed difference across proficiency level clusters in each group was due at least in part to differences in whether the students underwent the SPF process which provided them with dialogic interactive social interaction opportunities with their peers.

The fourth question examined the recurring themes in SPF as perceived and experienced by the participants in the experimental group. The participants ranked the aspects of writing in which SPF was perceived to be most effective as follows: global/macro aspects at 40% and local/micro aspects at 60%. Most of the participants in the focus group discussion (FGD) identified local/micro as the aspect in which SPF was most effective. As student respondent E002 articulated *“It helped me a lot because it helped me understand (the) rubric better. I*

mirrored my report on my classmates. We both discovered our common mistakes, especially during the peer conference.” E006 concurred by saying that “It helped me learn through someone’s work in terms of content, style, and structure because I saw the strengths in other works and gaps in my own”. However, E009 noted that “Although I learned how to catch my mistakes in my writing that I didn’t notice when I wrote it, I still prefer a professional to do the job.”

This observation led this current study to ask the participants to rank perceived conditions that made SPF the least effective. The following were identified: inaccuracies in feedback comments at 70% and inability to clarify and justify the feedback comments at 30%. E004 lamented that *“It was not very helpful because my partner did not know how to do the review. Some comments are not accurate and very basic.”* In addition, E003 expressed dismay, which was concurred by E001, over the same grievance that *“Some of the guide questions in the peer review form were not properly commented on by my pair and some were left blank, especially those that were related to content and advanced grammar aspects”*.

Finally, when asked about what should be done to improve the SPF process, the participants identified the following factors, in ranking order, as follow: comprehensive training at 60%, easy-to-use peer review form at 20%, and grade incentive at 20%. E007 clarified that *“Peer review is okay, but we need more training to do it on time”*. Raising another point, E0010 expressed that *“Students who do peer review well should be rewarded with some extra marks”*. Both E006 and E009 conjectured that *“Students could have done peer review better if the forms were easier to use for students.”* (Please see Appendix J).

Implications to Practice

The fifth and final question asked about the implications of SPF on ESL teaching. As the findings indicate, SPF benefits low-proficiency students under the following conditions: when they are paired with high-proficiency L2 students; when they are provided scaffolding support through peer review training, and when they are provided with the opportunity to interact with their peers in a peer review conference.

For SPF to be effective, it must be implemented as a process. The SPF process requires three phases: the pre-review phase which involves the design of the SPF review form and student training orientation of the form; the review phase which is comprised of two separate sessions on global and local aspects of writing; and the post-review phase which includes the peer conference.

When designing the SPF form, two requirements must be factored in: First, it must be based on the writing rubric which has a check listed local-focused feedback section and an open-ended global-focused feedback section given in two separate sessions (En-Chong, 2022). Second, it must be easy to use for students. SPF form should be able to support especially those less confident in their writing or their understanding of the assignment. The form is most helpful if they would be able to invite students to assess and describe elements of the writing rather than pass on quick judgments (Bean, 2011).

To keep low-proficiency students in the loop, it is suggested that you guide and monitor them during the course of the review process and the conference. Randomly check the implementation of review comments they received during revision. More scaffolding support is required for organization, cohesion, complex sentence writing, and correction of run-on

sentence errors which is deemed an L1 interference among Arabic speaking student-participants in this study. On the other hand, the main goals of the SPF training session must be concentrated on the following: to give local and global peer feedback based on the SPF form and to respond to peer feedback by developing sociolinguistic competence such as negotiation and mediation skills that include clarifying, justifying and confirming (Brammer & Calera, 2007; Chang, 2016).

Untrained student reviewers tend to focus on local issues such as spelling, grammar, vocabulary, and syntax. That rationalizes the need for training and to conduct two separate sessions exclusively focused on global and local aspects of writing. Having separate sessions ensures sufficient time to read and examine writing with one focus at a given time. The peer conference phase, alternatively, provides the opportunity for the peers to clarify, justify and confirm the feedback comments they have received and given to each other.

The whole SPF process requires an adequate amount of instructional time. The curriculum design in most ESL/EFL contexts may not be viable or sustainable for SPF practice. Nevertheless, if the curriculum itself could integrate SPF into the whole L2 writing process, particularly during formative assessments, then the issue of time constraints would be addressed proactively by mindful curriculum planning.

Therefore, SPF should be viewed as a pedagogical approach embedded in ESL/EFL curriculum that empowers students as their perspectives are exploited to supplement teachers' feedback. Hence, SPF is seen to scaffold social, linguistic, and cognitive support from peers to enable them to have multiple viewpoints in idea generation and development.

Teachers' feedback must reinforce important and accurate points to converge with those from student reviewers. In this manner, teachers emphasize the value of peer feedback that student reviewers can generate from the SPF process. Having said that, SPF should be considered as part of a graded assessment as well as testing critical thinking or evaluation skill to increase student motivation and keep them on task.

Conclusion

The findings of this current study indicate that SPF helped low-proficiency L2 students improve their writing proficiency. SPF supports them by making significant gains in both global and local aspects of their writing. Notable improvements are observed in the areas of content, grammar, spelling, vocabulary, and word choice. In addition, the findings also suggest that students demonstrated sociolinguistic competence in terms of their ability to clarify and respond to the feedback given.

The results of this study also may be explained in light of sociocultural theory. As seen, one benefit of the SPF process is the interaction between the L2 students of different proficiency levels. Specifically, high-proficiency students focus on and most likely provide feedback that is within their ZPD. Since their ZPDs are higher than that of their low-proficiency student-partners, the latter receives feedback that scaffolds their learning and benefits them in the process.

However, in the present study, such interaction occurs only between heterogenous pairs or high-low proficiency students. Hence, future research on the subject should examine SPF interaction impact on homogeneous pairs or high-high and low-low proficiency students as well. Nevertheless, the findings are relevant since they suggest that SPF may be even more

beneficial to low-proficiency students than earlier thought. Thus, to investigate further the findings of this study, future studies should examine the effects of the SPF process on higher language proficiency levels at C1 and C2 levels, on L2 writing that covers other genres such as narrative and argumentative essays, and on L1 students.

Finally, although making the whole SPF process effective requires sufficient time in planning the review, designing the form, and most importantly training students, the whole review process itself works especially in scaffolding students with sociolinguistic skills that can enable them to become more proficient writers.

About the Author

Carlo Domingo Casinto is an assistant professor at the School of Foundation at Bahrain Polytechnic. His current research interests include peer feedback L2 academic writing, virtual remote teaching pedagogy, CEFR framework adoption for ESL/EFL curricula, and reflective teaching. ORCID ID: 0000-0002-2103-5873.

To Cite this Article

Casinto, C. D. (2023). Scaffolded peer feedback and second language writing proficiency: Implications for inclusive ESL teaching practice. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, 26 (4). <https://doi.org/10.55593/ej.26104a8>

References

- Alnasser, S.M.N. (2018). Employment of peer feedback technique in L2 writing classrooms: An introductory guide for novice instructors. *Universal Journal of Educational Research*, 6 (8), 1652-1658. <http://doi.org/10.13189/ujer.2018.060804>
- Alsehibany, R.A. (2021). EFL Saudi students' attitudes toward peer feedback activities in a writing class. *PSU Research Review*, 1747-2399. <http://doi.org/10.1108/PRR-01-2021-0004>
- Bagheri, M. & Rassaei, E. (2022). The effects of two forms of written corrective feedback and ambiguity tolerance on EFL learners' writing accuracy. *English Teaching & Learning*, 46, 19–38. <https://doi.org/10.1007/s42321-021-00082-6>
- Bean, J.C. (2011). *Engaging ideas: A professor's guide to integrating writing, critical thinking, and active learning in the classroom (2nd ed)*. Jossey-Bass.
- Bolourchi, A. & Soleimani, M. (2021). The impact of peer feedback on EFL learners' writing performance and writing anxiety. *International Journal of Research in English Education*, 6(1), 1-15. <https://ijreeonline.com/article-1-425-en.pdf>
- Brammer, C. & Calera, M.R. (2007). Peer review from students' perspective: Invaluable or Invalid? *Composition Studies*, 35(2). 71-85. <https://www.jstor.org/stable/43501704>
- Chang, C.Y. (2016). Two decades of research in L2 peer review. *Journal of Writing Research*, 8(1), 81-117. <http://doi.org/10.17239/jowr-2016.08.01.03>
- En-Chong, L. (2022). Effects of Rubric-Based and Detailed Peer Feedback on University-Level English as a Foreign Language Students' Writing Self-efficacy and Subsequent Revisions. *English Teaching & Learning*, 46, 59–76. <https://doi.org/10.1007/s42321-021-00088-0>

- Do, T.B.T. (2020). Effects of scaffolded peer review training on the quality produced by students of French as a foreign language. *Electronic Journal of Foreign Language Teaching*, 7(1), 318-300. <https://e-flt.nus.edu.sg/wp-content/uploads/2020/09/do.pdf>
- Fatoni, M. and Tauchid, A. (2020). An analysis of guided peer review to improve writing skill on personal recount texts for high school students in Bojonegoro. *Jurnal Pendidikan Edutama*, 7(1), 63-74. <http://ejurnal.ikipgribojonegoro.ac.id/index.php/JPE>
- Hislop, J. & Stracke, E. (2017). ESL students in peer review: An action research study in a university English for Academic Purposes course. *University of Sydney Papers in TESOL*, 12, 9-44. http://faculty.edfac.usyd.edu.au/projects/usp_in_tesol/pdf/volume12/Article02.pdf
- Hossein, N. & Eva, K. (Eds.) (2017). *Corrective Feedback in Second Language Teaching and Learning: Research, Theory, Applications, Implications (1st ed.)*. Routledge. <https://doi.org/10.4324/9781315621432>
- Hu, G. & Lam, S. T. E. (2010). Issues of cultural appropriateness and pedagogical efficacy: exploring peer review in a second language writing class. *Instructional Science*, 38(4), 371–394. <http://www.jstor.org/stable/23372855>
- Kamimura, T. (2006). Effects of peer review on EFL student writers at different levels of English proficiency: A Japanese context. *TESL Canada Journal*, 23(2), 2-39.
- Kelly, L. (2015). Effectiveness of Guided Peer Review of student essays in a large undergraduate Biology course. *International Journal of Teaching and Learning in Higher Education*. 27(1), 56-68. <https://files.eric.ed.gov/fulltext/EJ1069821.pdf>
- Kim, Y. & Emeliyanova, L. (2021). The effects of written corrective feedback on the accuracy of L2 writing: comparing collaborative and individual revision behavior. *Language Teaching Research*, 25(2), 234-255. <https://teslcanadajournal.ca/index.php/tesl/article/view/53/53>
- Lantolf, J. & Beckett, T. (2009). Sociocultural theory and second language acquisition. *Language Teaching*, 42(4), 459-475. <http://doi.org/0.1017/S0261444809990048>
- Leijen, D. A. J. (2017). A novel approach to examine the impact of web-based peer review on the revisions of L2 writers. *Computers and Composition*, 53, 35-54. <https://doi.org/10.1016/j.compcom.2016.11.005>
- Lou, Y., Li, Z., Gong, P. & Liu, J. (2016). A study on guided peer feedback in group work to improve non-English-majored graduates' English writing in internet-based language laboratory", *Open Journal of Sciences*, 4, 86-96. <https://www.scirp.org/journal/paperinformation.aspx?paperid=71389>
- Lundstrom, K. & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's writing. *Journal of Second Language Writing*, 18, 30-43. <https://doi.org/10.1016/j.jslw.2008.06.002>
- Min, H.T. & Chiu, Y. M. (2021). The relative effects of giving versus receiving comments on students' revision in an EFL writing class. *English Teaching & Learning*. <https://doi.org/10.1007/s42321-021-00094-2>

- Nassaji, H. (2016). Interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research*, 20 (4), 535-562. <http://doi.org/10.1177/1362168816644940>
- Nguyen, C.D. (2021). Scaffolding student engagement with written corrective feedback: transforming feedback sessions into learning affordances. *Language Teaching Research*. 1-22. <https://doi.org/0.1177/13621688211040904>
- Saeli, H. & Cheng, A. (2021). Peer feedback, learners' engagement, and L2 writing development: The case of a test-preparation class. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, 25 (2), <https://tesl-ej.org/pdf/ej98/a7.pdf>
- Salih, A.R.A. (2013). Peer response to L2 student writing: patterns and expectations. *English Language Teaching*, 6(3), <https://doi.org/10.5539/elt.v6n3p42>
- Sanchez-Naranjo, J. (2019). Peer review and training: Pathways to quality and value in second language writing. *Foreign Language Annals*, 52(3), 612-643. <https://doi.org/10.1111/flan.12414>
- Sippel, L. (2021). Maximizing the benefits of peer interaction: form-focused instruction and peer feedback training. *Language Teaching Research*. 1-27, <https://doi.org/10.1177/13621688211004638>
- So, L. & Lee, C. H. (2012). Peer feedback using blended learning in L2 writing at the university level. *English Teaching*, 67(3), 307-337. http://journal.kate.or.kr/wp-content/uploads/2015/01/kate_67_3_13.pdf
- Sotouddehnama, E. & Pilehvari, A. (2016). The impact of peer review on EFL learners' writing proficiency: global and local aspects. *Porta Linguarum*, 25, 35-47. <https://dialnet.unirioja.es/descarga/articulo/5412394.pdf>
- Storch, N. & Aldosari, A. (2012). Pairing learners in pair work activity. *Language Teaching Research*, 17(1), 31-48. <https://doi.org/10.1177/1362168812457530>
- Storch, N & Suzuki, W. (2020). *Languaging in language learning and teaching: A collection of empirical studies*. John Benjamins Publishing Company. <https://doi.org/10.1075/llt.55>
- Tajabadi, A., Ahmadian, M., Dowlatabadi, H. & Yazdani, H. (2020). EFL learners' peer negotiated feedback, revision outcomes, and short-term writing development: The effects of patterns of interaction. *Language Teaching Research*. 1-29, <https://doi.org/10.1177/1362168820951207>
- Uymaz, E. (2019). The effects of peer feedback on the essay writing performances of EFL students. *International Journal of Curriculum and Instruction*, 11(2), 20-37.
- Vygotsky, L.S. (1986). *Thought and Language*. MIT Press.
- Xie, J. & Mu, F. (2015). A Study into Students' Views on Guided Peer Feedback in Group Work. *Chinese Journal of Applied Linguistics*, 38(2), 188-203. <https://doi.org/10.1515/cjal-2015-0011>
- Zhao, H. (2018). New insights into the process of peer review for EFL writing: a process-oriented socio-cultural perspective. *Learning and Instruction*, 58, 263-273. <https://doi.org/10.1016/j.learninstruct.2018.07.010>

Zheng, Z. (2012). Understanding the learning process of peer feedback activity: An ethnographic study of exploratory practice. *Language Teaching Research*, 16 (1), 109-126. <https://doi.org/10.1177/1362168811426248>

Zhu, W. (2001). Interaction and feedback in mixed peer response groups. *Journal of Second Language Writing*, 10, 251-276. [https://doi.org/10.1016/S1060-3743\(01\)00043-1](https://doi.org/10.1016/S1060-3743(01)00043-1)

APPENDICES

Appendix A

Post-test on Listening & Writing (LW) Assessment

Date:	Week 15 (Tuesday 15 th June 2021)
Time:	11:45am – 02:15pm (two hours and a half)
Grade value:	40%
Submission:	Electronic (via Moodle)
Assessment type:	Individual
Word count:	300 – 350 words

Learning Outcomes

1. Demonstrate a range of routine and non-routine skills to obtain, process and respond to a variety of written inputs and produce a range of written texts to draw conclusions and suggest solutions in defined situations.
2. Demonstrate a range of routine and non-routine skills to obtain, process and respond to a variety of oral/aural inputs and produce a range of communicative responses to draw conclusions and suggest solutions in defined situations.

Note:

- To pass this course, a student needs to pass learning outcomes 1 and 2.
- This assessment assesses both learning outcomes 1 and 2. **To pass this the course**, a student must achieve an aggregate course grade of 60/100 or above.

Instructions

- This assessment consists of two parts: listening and writing. You are expected to complete **both** parts.
- You have **two hours and a half** to complete this assessment (start 11:45am, finish 02:15pm)
- Save your work before uploading it to Moodle.
- Contact your tutor through BBB or email in case you have any technical issues.

Scenario

Bahrain Polytechnic is planning to send their student football team to either **Brazil** or **Argentina** for remote football training sessions. They also want the team to learn the Spanish language. You are a member of Bahrain Polytechnic Student Council (BPSC). The council has assigned you to write a mini proposal to recommend one of these countries to send the team to.

Listening

- Play the audio recording **at least twice** (but you have unlimited access to it).
- Listen to the audio comparing and contrasting Brazil and Argentina in terms of **demographics** and **culture**.

While you listen, make relevant notes in the table below. Some of the notes have been already made for you.

		Brazil	Argentina
Demographics	Historical origin (Similarity)		
	Natural resources (Similarity)		
	Colonization (Difference)		
	Population (Difference)		
Culture	People (Similarity)		
	Sport (Similarity)		
	Main language (Difference)		
	Food (Difference)		

Writing

- Open a new **MS word document**.
- Save it as **EL5002_Summative_LW**.
- **DO NOT** add your name or ID to the file name.
- Use **Calibri** font, **size 11** with **1.5 line spacing**.
- Use your notes from the Listening section to write a mini proposal (**300—350 words**).
- Remember to paraphrase to avoid plagiarism.
- Use this format for your proposal:
 - **Introduction:** include background information, a two-options statement, and a thesis statement.
 - **Main body** (must be broken down into 2 paragraphs with headings): include a topic sentence, comparative language/compare and contrast sentences and a concluding sentence saying which option is better – you must include at least 1 similarity and 2 differences in each body paragraph.
 - **Recommended course of actions:** include your recommendation (claim), support it with evidence and reasoning and mention the next step (next course of action).

File submission

- Submit only your proposal **without** the table that you used for writing the notes.
- Before you submit, please check you have done the following:
 - Introduction:**
 - background information (scenario)
 - 2 options statement
 - thesis statement
 - Body Paragraphs 1 and 2**
 - topic sentence
 - 1 similarity
 - 2 differences
 - concluding sentence
 - Recommendation:**
 - CER
 - Next course of action
- Upload your proposal to Moodle **Assessments/Listening and Writing Assessment (40%)**

End of Assessment

Appendix B

Post -test on Listening and Writing (LW) Assessment Rubric¹

Task fulfilment	4	8	12	16
	<ul style="list-style-type: none"> • covers 2 points or less from the listening • no attempt at paraphrasing 	<ul style="list-style-type: none"> • elements of 4 and 12 	<ul style="list-style-type: none"> • covers 4 points from the listening • demonstrates some ability to paraphrase 	<ul style="list-style-type: none"> • elements of 1
	2	4	6	8
Grammar²	<ul style="list-style-type: none"> • most grammatical structures are inappropriate and meaning is distorted and obscure 	<ul style="list-style-type: none"> • elements of two and six 	<ul style="list-style-type: none"> • uses a limited range of appropriate grammatical structures with frequent grammatical errors which may cause slight difficulty for the reader 	<ul style="list-style-type: none"> • elements of s ten

¹ students must meet **all** requirements in the descriptor to be awarded a particular score

² B1/B2 grammar (language of compare and contrast/compound and complex sentences)

Organisation & Cohesion, Coherence	<ul style="list-style-type: none"> • does not meet format requirements (headings, font, spacing) • relationships between ideas and content missing or faulty to the point of causing significant strain for the reader • inappropriate introduction, body paragraphs & recommended course of action • faulty cohesive devices (repetitive, inaccurate) causes confusion for the reader • lacking referencing 	<ul style="list-style-type: none"> • elements of two and six 	<ul style="list-style-type: none"> • meets some format requirements (headings, font, spacing) • adequate introduction, body paragraphs & recommended course of action • relationships between ideas and content indicated • range of cohesive devices used • limited or repeated referencing 	<ul style="list-style-type: none"> • elements of ten
	1	2	3	4
CERA⁵ (claim evidence, reasoning, action)	<ul style="list-style-type: none"> • no relevant claim given with no evidence and/or reasoning in the Recommended Course of Action paragraph 	<ul style="list-style-type: none"> • elements of one and three 	<ul style="list-style-type: none"> • relevant claim given and supported with little evidence and reasoning in the Recommended Course of Action paragraph 	<ul style="list-style-type: none"> • elements of five
Vocabulary range + accuracy	<ul style="list-style-type: none"> • basic vocabulary mostly used inappropriately with limited control which impedes communication 	<ul style="list-style-type: none"> • elements of one and three 	<ul style="list-style-type: none"> • adequate range used appropriately with some inaccuracies which do not impede communication 	<ul style="list-style-type: none"> • elements of five

³ linking words, phrases or clauses to connect ideas and information

⁴ use of a word or phrase (such as this, that, these, those, now, then) that points to the time, place, or situation in which the speaker is speaking

⁵ 'claims' are the writer's main points, statements, or ideas that need to be backed up by explanation and examples/evidence/action (order can vary)

Register	<ul style="list-style-type: none"> • inappropriate tone • slang and/or idiomatic expressions • 3 or more contractions 	<ul style="list-style-type: none"> • tone varies • minor slips in use of slang or idiomatic expressions • 2 or less contractions 	<ul style="list-style-type: none"> • tone appropriate to the audience and purpose consistently (throughout) • no inappropriate use of slang or idiomatic expressions • no contractions 	
Wordcount	<ul style="list-style-type: none"> • 269 or less; 371 or more words 	<ul style="list-style-type: none"> • 270-299 or 351-370 words 	<ul style="list-style-type: none"> • within word count limit: 300-350 words 	
Punctuation (including capitalization)	<ul style="list-style-type: none"> • 3 or more errors in punctuation 	<ul style="list-style-type: none"> • 2 errors in punctuation 	<ul style="list-style-type: none"> • 1 or no errors in punctuation 	
Spelling	<ul style="list-style-type: none"> • more than 1 spelling mistake 	<ul style="list-style-type: none"> • 1 or no spelling mistake 		

Appendix C
Peer Review Form

Writer's Name: _____

Reviewer's Name: _____

Part I: CHECKLIST (Duration: 40 minutes)

Read the draft and look for evidence to answer the questions with a YES or a NO.

✓ **A. Content Check**

Headings	Parts	Question	YES	NO
INTRODUCTION	Background	Is there a <u>general and specific reasons</u> why the proposal report is written?		
	Options	Are there <u>two options</u> presented for comparison & contrast?		
	Purpose	Is there an <u>explanation on the purpose of the report and basis of the recommendation</u> ?		
BODY 1st	Topic Sentence	Is there an <u>explanation on the importance of the factor</u> that the options are compared and contrasted on?		
	Supporting Sentence	Is there <u>one similarity and two differences</u> of the options?		
	Concluding Sentence	Is there a <u>decision</u> as to which option is better?		
	Topic Sentence	Is there an <u>explanation on the importance of the factor</u> that the options are compared and contrasted on?		
	Supporting Sentences	Is there <u>one similarity and two differences</u> of the options?		
	Concluding Sentence	Is there a <u>decision</u> as to which option is better?		
CONCLUSION	Claim	Is there an <u>overall choice/recommendation</u> between two options?		
	Evidence	Is there a <u>summary of evidence</u> presented in the body paragraphs?		
	Reason	Is there a <u>summary of rationale/reason</u> of importance of evidence presented in the body paragraphs?		
	Next Course of Action	Is there a <u>recommended next logical action (s)</u> to do?		

✓ **B. Coherence, Grammar, Tone/Register and Word Count**

Rubric Criteria	Question	YES	NO	If YES, which paragraph & sentence?
Coherence	Are there missing <u>appropriate linking words</u> for comparison, contrast and conclusion?			
Grammar & Punctuation	Are there <u>errors on subject-verb agreement, comparative adjectives, run-on</u> or missing commas and full stops?			
Tone & Register	Are there <u>errors on contractions and use of 1st & 2nd person personal pronouns</u> ?			
Word Count	Is the <u>report within 300-350 words</u> ?			

Part II. EVALUATION FEEDBACK (Duration: 20 minutes)

Read the draft again and answer the following questions in one or two sentences.

1) Is the thesis statement or purpose of the report clear? Why or why not?

2) Is the evidence presented and discussed in the body paragraphs convincing? Why or why not?

3) Do you agree with the recommended next course of action? Why or why not?

4) In a scale of 1(lowest)-10 (highest), how would you rate this report? Why?

5) What do you like the most in this report (content details, coherency, grammar, etc.)?

6) What should this report need to improve on (content details, coherency, grammar, etc.)?

Instructor's Note:

Bring this completed Peer Review Form to your Peer Review Conference with the writer. The Peer Review Conference shall be held on the second hour of the Peer Review Session.

-----END OF PEER REVIEW-----

Appendix D

Peer Review Focus Group Discussion (FGD) Guide

FGD GUIDE QUESTIONS for Scaffolded Peer Feedback Sessions

Part I. During the Peer Review

- 1) By looking into your classmate's draft, did you find PEER REVIEW helpful in reflecting on your report as well? How?
- 2) What is it in PEER REVIEW that you don't like?
- 3) What do you think should be done to make PEER REVIEW more effective?
- 4) As it currently is, in a scale of 1(lowest)-10 (highest), how do rate PEER REVIEW as a technique in writing? Why?

Part II. During the Peer Review Conference

- 1) By discussing with your classmate his/her draft, did you find PEER REVIEW helpful in reflecting on your report as well? How?
- 2) What is it in Peer Review Conference that you don't like?
- 3) What do you think should be done to make Peer Review Conference more effective?
- 4) As it currently is, in a scale of 1(lowest)-10 (highest), how do rate Peer Review Conference as part of the Peer Review Process? Why?

Appendix E

Means and Gain Scores of Experimental Group

<i>Student Name</i>	<i>ID Number</i>	<i>Code Number</i>	<i>CEFR Level</i>	<i>AVERAGE FORMATIVE SCORE</i>	<i>SUMMATIVE SCORE</i>	<i>GAIN SCORE</i>
		E003	B2	77	88.52	11.52
		E001	B2	81.96	83.61	1.65
		E002	B1	55.7	78.69	22.99
		E006	B1	81.9	96.72	14.8
		E004	B2	93	96.72	3.72
		E008	B2	81.9	88.52	6.6
		E009	B1	85.24	95.08	9.84
		E007	B2	78.6	86.89	8.2
		E010	B1	86.88	98.36	11.48
		E005	B1	62.2	80.33	18.04
Group Mean Score	v			78.43	89.34	10.91%
B1 Mean Score				74.38	88.19	13.81%
B2 Mean Score				82.49	88.79	6.30%

Appendix F

Means and Gain Scores of Control Group

<i>Student Name</i>	<i>ID Number</i>	<i>Code Number</i>	<i>CEFR Level</i>	<i>AVERAGE FORMATIVE SCORE</i>	<i>SUMMATIVE SCORE</i>	<i>GAIN SCORE</i>
		C001	B2	90.16	91.80	1.63%
		C005	B1	65.57	67.21	1.63%
		C004	B1	67.2	85.25	18.05%
		C003	B2	83.60	88.52	4.92%
		C002	B1	73.77	86.89	13.11%
		C008	B2	83.60	86.89	3.28%
		C009	B2	86.88	91.80	4.91%
		C006	B1	78.68	86.89	8.20%
		C007	B2	81.9	90.16	8.19%
		C010	B1	77	81.97	4.97%
Group Mean Score				78.83	85.73	6.90%
B1 Mean Score				72.44	81.64	9.20%
B2 Mean Score				85.22	89.83	4.61%

Appendix G:

<i>Student Name</i>	<i>ID Number</i>	<i>Code Number</i>	<i>CEFR Level</i>	<i>Formative Errors</i>	<i>Summative Errors</i>	<i>Error Avoidance Efficiency Rate</i>
		E003	B2	E02	E02	0%
		E001	B2	E02; E07	E02	50%
		E002	B1	E01; E02; E03; E04; E05; E06; E07	E02; E03; E05;	42.85%
		E006	B2	E02; E07	E02	50%
		E004	B2	E02	E02	0%
		E008	B1	E01; E02; E03; E06	E02; E03	0%
		E009	B1	E02; E06	E02	50%
		E007	B2	E02; E03; E04; E06	E02; E03; E06	25%
		E010	B1	E02	E02	0%
		E005	B1	E01; E02; E03; E05; E06; E07	E02; E03; E07	20%
Group Mean Rate						23.78% Higher than Control Group
B1 Mean Rate						22.57% Lower than B2
B2 Mean Rate						25% Higher than B1

Common Errors Avoidance Efficiency Rate of Experimental Group

Error Codes:

Code	Description
E01	Content
E02	Grammar (including punctuation & capitalization)
E03	Organization & Cohesion
E04	Spelling
E05	CERA (claim, evidence, reason, action)
E06	Vocabulary (including register)
E07	Word Count

Appendix H:

<i>Student Name</i>	<i>ID Number</i>	<i>Code Number</i>	<i>CEFR Level</i>	<i>Formative Errors</i>	<i>Summative Errors</i>	<i>Error Avoidance Efficiency Rate</i>
		C001	B2	E02; E06	E02	50%
		C005	B1	E01; E02; E04; E07	EQ1; EQ2; EQ4; E06; E07	-25%
		C004	B1	E02; E03	E02; E03	0%
		C003	B2	E02; E06	E02; E06	0%
		C002	B1	E02; E04; E06	E02; E06	33.33%
		C008	B2	E02; E03	E02; E03	0%
		C009	B2	E02; E06	E02	50%
		C006	B1	E01; E02; E03; E06	E01; E02; E03; E06	0%
		C007	B2	E02	E02	0%
		C010	B1	E02; E04; E06	E02; E04; E06	0%
Group Mean Rate						13.33% Lower than Experimental Group
B1 Mean Rate						6.66% Lower than B2
B2 Mean Rate						20% Higher than B1

Common Errors Avoidance Efficiency Rate of Control Group

Error Codes

Code	Description
E01	Content
E02	Grammar (including punctuation & capitalization)
E03	Organization & Cohesion
E04	Spelling
E05	CERA (claim, evidence, reason, action)
E06	Vocabulary (including register)
E07	Word Count

Appendix I

Data Sets

By Group

t-Test: Paired Two Sample for Means		
	<i>Control GRP</i>	<i>Experimental GRP</i>
Mean	85.738	89.344
Variance	51.38815111	50.89104889
Observations	10	10
Pearson Correlation	0.522084034	
Hypothesized Mean	0	
df	9	
t Stat	-1.63099823	
P(T<=t) one-tail	0.068663842	
t Critical one-tail	1.833112933	
P(T<=t) two-tail	0.137327684	
t Critical two-tail	2.262157163	

<i>Control GRP</i>		<i>Experimental GRP</i>	
Mean	85.738	Mean	89.344
Standard Error	2.266895	Standard Error	2.255904
Median	86.89	Median	88.52
Mode	86.89	Mode	88.52
Standard Deviation	7.168553	Standard Deviation	7.133796
Sample Variance	51.38815	Sample Variance	50.89105
Kurtosis	5.677393	Kurtosis	-1.456
Skewness	-2.20784	Skewness	-0.16142
Range	24.59	Range	19.67
Minimum	67.21	Minimum	78.69
Maximum	91.8	Maximum	98.36
Sum	857.38	Sum	893.44
Count	10	Count	10

By B1 Proficiency Level

t-Test: Paired Two Sample for Means		
	<i>B1 Control</i>	<i>B1 Experimental</i>
Mean	81.642	89.836
Variance	69.12272	90.53623
Observations	5	5
Pearson Correlati	0.80533047	
Hypothesized Me	0	
df	4	
t Stat	-3.22676308	
P(T<=t) one-tail	0.01603694	
t Critical one-tail	2.13184679	
P(T<=t) two-tail	0.03207388	
t Critical two-tail	2.77644511	

<i>B1 Control</i>		<i>B1 Experimental</i>	
Mean	81.642	Mean	89.836
Standard Error	3.718137	Standard Error	4.255261
Median	85.25	Median	95.08
Mode	86.89	Mode	#N/A
Standard Deviation	8.314007	Standard Deviation	9.515053
Sample Variance	69.12272	Sample Variance	90.53623
Kurtosis	3.790671	Kurtosis	-3.13823
Skewness	-1.93573	Skewness	-0.55813
Range	19.68	Range	19.67
Minimum	67.21	Minimum	78.69
Maximum	86.89	Maximum	98.36
Sum	408.21	Sum	449.18
Count	5	Count	5

By B2 Proficiency Level

t-Test: Paired Two Sample for Means		
	<i>B2 Control</i>	<i>B2 Experimental</i>
Mean	89.834	88.852
Variance	4.55758	23.36347
Observations	5	5
Pearson Correlation	-0.44143	
Hypothesized Mean	0	
df	4	
t Stat	0.360838	
P(T<=t) one-tail	0.368235	
t Critical one-tail	2.131847	
P(T<=t) two-tail	0.73647	
t Critical two-tail	2.776445	

	<i>B2 Control</i>		<i>B2 Experimental</i>
Mean	89.834	Mean	88.852
Standard Error	0.954733	Standard Error	2.161642
Median	90.16	Median	88.52
Mode	91.8	Mode	88.52
Standard Deviation	2.134849	Standard Deviation	4.833577
Sample Variance	4.55758	Sample Variance	23.36347
Kurtosis	-1.50129	Kurtosis	2.539024
Skewness	-0.53754	Skewness	1.237514
Range	4.91	Range	13.11
Minimum	86.89	Minimum	83.61
Maximum	91.8	Maximum	96.72
Sum	449.17	Sum	444.26
Count	5	Count	5

Appendix J: Coded FGD Transcript

<i>Student Name</i>	<i>ID Number</i>	<i>Code Number</i>	<i>CEFR Level</i>	<i>Theme 1: Most effective</i>	<i>Theme 2: Least effective</i>	<i>Theme 3: Factor for improvement</i>
		E003	B2	Local/micro; working with a partner is quite a good idea especially for grammar check.	Inaccuracies in feedback; some of the guide questions in the peer review form were not properly commented on by my pair and some were left blank especially those that were related to content and advanced grammar aspects.	Comprehensive training; proper training will make us better in this task.
		E001	B2	Global/macro; generally helpful as you have another person's eyes look into your errors as you would have overlooked.	Inability to negotiate feedback; I agree with Sarah. Some questions were ignored.	Comprehensive training; indeed, more training will help.
		E002	B1	Local/micro; it helped me a lot because it helped me understand rubric better. I mirrored my report on my classmate's. We both discovered our common mistakes especially during the peer conference.	Inaccuracies in feedback; looking for evidence is difficult	Comprehensive training; I need to be taught how and where to look for errors.
		E006	B2	Global/macro; it helped me learned through someone's work in terms of content, style and structure because I saw the strengths in other works and gaps in my own.	Inability to negotiate feedback; the form was not completed well by my partner.	Easy-to-use form; make forms student friendly
		E004	B2	Global/macro; exchanging ideas on content development	Inability to negotiate feedback; it was not very helpful because my partner did not know how to do the review. Some comments are not accurate and very basic.	Grade incentive; this (peer review) should be part of graded assessments
		E008	B1	Local/micro; it was helpful in helping me figure out my mistakes.	Inaccuracies in feedback; took a lot of time to do yet some errors were not identified.	Comprehensive training; train us more.
201900805		E009	B1	Local/micro; although I learned how to catch my mistakes in my writing that I didn't notice when I wrote it, I still prefer a professional to do the job.	Inaccuracies in feedback; I'm not confident doing it. That's why I wasn't able to answer all review questions.	Easy-to-use form; students could have done peer review better if the forms were easier to use for students.
201902444		E007	B2	Global/macro; generally helpful for missing details and in organizational lapses.	Inaccuracies in feedback; some comments are not supported with evidence.	Comprehensive training; Peer review is okay, but we need more training to do it on time.
202001769		E010	B1	Local/micro; my classmate supported me in both content and grammar.	Inaccuracies in feedback; I'm not confident doing it for someone.	Grade incentive; students who do peer review well should be rewarded with some extra marks
202000695		E005	B1	Local/micro; it gave me a chance to see my classmate's work and learned something.	Inaccuracies in feedback; I need a much longer time to find out what the mistakes are.	Comprehensive training; I think (I) need more trainings.

<i>Student Name</i>	<i>ID Number</i>	<i>Code Number</i>	<i>CEFR Level</i>	<i>Theme 1: Most effective</i>	<i>Theme 2: Least effective</i>	<i>Theme 3: Factor for improvement</i>
		E003	B2	Local/micro; working with a partner is quite a good idea especially for grammar check.	Inaccuracies in feedback; some of the guide questions in the peer review form were not properly commented on by my pair and some were left blank especially those that were related to content and advanced grammar aspects.	Comprehensive training; proper training will make us better in this task.
		E001	B2	Global/macro; generally helpful as you have another person's eyes look into your errors as you would have overlooked.	Inability to negotiate feedback; I agree with Sarah. Some questions were ignored.	Comprehensive training; indeed, more training will help.
		E002	B1	Local/micro; it helped me a lot because it helped me understand rubric better. I mirrored my report on my classmate's. We both discovered our common mistakes especially during the peer conference.	Inaccuracies in feedback; looking for evidence is difficult	Comprehensive training; I need to be taught how and where to look for errors.
		E006	B2	Global/macro; it helped me learned through someone's work in terms of content, style and structure because I saw the strengths in other works and gaps in my own.	Inability to negotiate feedback; the form was not completed well by my partner.	Easy-to-use form; make forms student friendly
		E004	B2	Global/macro; exchanging ideas on content development	Inability to negotiate feedback; it was not very helpful because my partner did not know how to do the review. Some comments are not accurate and very basic.	Grade incentive; this (peer review) should be part of graded assessments
		E008	B1	Local/micro; it was helpful in helping me figure out my mistakes.	Inaccuracies in feedback; took a lot of time to do yet some errors were not identified.	Comprehensive training; train us more.
		E009	B1	Local/micro; although I learned how to catch my mistakes in my writing that I didn't notice when I wrote it, I still prefer a professional to do the job.	Inaccuracies in feedback; I'm not confident doing it. That's why I wasn't able to answer all review questions.	Easy-to-use form; students could have done peer review better if the forms were easier to use for students.
		E007	B2	Global/macro; generally helpful for missing details and in organizational lapses.	Inaccuracies in feedback; some comments are not supported with evidence.	Comprehensive training; Peer review is okay, but we need more training to do it on time.
		E010	B1	Local/micro; my classmate supported me in both content and grammar.	Inaccuracies in feedback; I'm not confident doing it for someone.	Grade incentive; students who do peer review well should be rewarded with some extra marks
		E005	B1	Local/micro; it gave me a chance to see my classmate's work and learned something.	Inaccuracies in feedback; I need a much longer time to find out what the mistakes are.	Comprehensive training; I think (I) need more trainings.

Copyright of articles rests with the authors. Please cite TESL-EJ appropriately.