Effects of Collaborative Evaluation Techniques on Students’ Writing Performance

ALMA CITA B. SORIA-CALIMBO*
MAYFLOR D. PRANTILLA-ARAMBALA
AMITA B. SUSUSCO
Languages and Literature, Central Mindanao University
Philippines

ONNAH PIERRE P. TALLE
English Department, Mindanao State University-Iligan Institute of Technology
Philippines

ABSTRACT

The study utilized a pre-test/post-test quasi-experimental design to investigate the effects of four evaluation techniques (self, one-on-one, and two peer-group editing techniques) on students’ writing performance. Participants were 88 college students enrolled in three sections of freshmen writing taught by one teacher. All student-editors were guided by an adapted holistic rubric considering three features of the compositions, namely: content (focus and support), coherence (organization) and accuracy (grammar and mechanics). Writing performance was measured through the average holistic scores of students in the post-test in comparison with their scores in the pre-test. Statistical tests included ANOVA and the Tukey HSD procedure. Results showed that the evaluation methods had a significant impact on all of the components of the rubric. With an inter-rater reliability of .80, the average difference in the holistic scores between post-test and pre-test was statistically significant. It was also found that except for mechanics, the average scores for content and organization in the essays subjected to collaborative evaluation techniques (one-on-one and peer group) were significantly higher than those subjected to the self-editing technique. It is strongly recommended that collaborative evaluation be an integral part in any writing instruction.

KEYWORDS: collaborative learning, editing techniques, teaching writing

*corresponding author
Introduction

Studies show that there is a decline in students’ writing quality (Pondiscio, 2010; Lingwall, 2010; Dieterich, 2010). Too little writing practice and too little feedback are among the reasons for this sad reality (Blasingame & Bushman, 2005; Karegianes, 1980). Students’ frequent practice and teachers’ immediate feedback are indeed crucial in any learning process (Erin, 2010). But giving feedback to students’ written products is the most challenging part in teaching writing (Alcantara et al., 2003; Hyslop, 2010). An average of three writing activities, for instance, can make giving immediate feedback impossible for the teacher who is going to evaluate 150 essays in a week. Due to much time required in reading and evaluating students’ compositions, teachers reasonably give students lesser writing tasks which they can manage to evaluate; thus, in a sense depriving the students of the necessary practice needed to enhance their writing skills.

One solution to such a problem is to use peers in evaluating students’ writing. With peer evaluation, both teachers and students stand to benefit. It reduces the workload of writing instructors giving them more time in designing a good number of writing tasks rather than in checking a huge number of individual compositions. Students, on the other hand, will be actively involved in the learning process (Blasingame & Bushman, 2005), will develop evaluation skills and become more responsible for their own learning (Haaga, as cited by Cho, Schunn, & Wilson, 2006). However, collaborative evaluation of writing outputs, if not properly mediated by teachers, can be a waste of time.

Since giving feedback to student writing is the most challenging part in teaching writing, a number of studies have been conducted to find out how to make the task a meaningful one. Responding to student writing is basically done by the writing instructor; however, despite the fact that it is so time-consuming and so tedious on the part of the teacher, it is unrewarding in the sense that teacher correction, whether detailed or less detailed, does not guarantee students’ understanding of the responses which made them simply ignore these. Another thing, teacher correction affects students’ confidence and motivation as they suffer “death by the red pen” (Zheng, 2007, p.26).

Since writing is basically done solo, stages in writing are also basically done solo, including doing self-editing. At the college level, assessment commonly includes shorter essay writing which can be subjected to self-editing before submitting the final paper. However, Kasule and Lunga (2010) investigated the effect of self-editing on students’ writing tasks and found the failure of the method in minimizing students’ errors especially in sentence construction and organization of ideas. It was then recommended that other evaluation methods, namely by peers or the computer may be given consideration.

Collaborative evaluation of writing outputs is basically anchored on the principle of collaboration in learning which “involves students working in teams to accomplish a common goal” (Felder & Brent, 2007, p.2). Effective instruction encourages students to work collaboratively in groups to make learning more engaging and more meaningful. According to Vygotsky (in Dooley, 2008, p.2), collaborative learning situations will enable students to “perform at higher intellectual levels.” Gokhale (1995) confirmed this when he investigated whether there is a significant difference in the performance between students who work in groups and those who work individually, and concluded that both methods are effective in gaining factual knowledge but for higher order thinking skills, collaborative learning is more effective. In another study, it was also found that collaborative peer editing helps students develop critical thinking and editorial skills and promotes teamwork (Nelson, 2002). Since giving feedback to writing requires much
intellectual activity, it can be effectively done collaboratively (Hyslop, 2010). Both teachers and students acknowledge its importance in writing instruction (Balushy, n.d.)

A considerable body of literature shows that peer editing as a mode of giving feedback to writing has been found an effective tool in teaching writing proficiency. Goodman et al. (n.d.) found that peer review and editing improved the quality of medical reporting. Investigating the effects of peer versus teacher editing on writing achievement, Karegianes (1980) concluded that the peer edit group had significantly higher writing proficiency than the teacher edit group. Berg (1997) investigated the effects of trained peer response on students’ writing outcomes and found that trained peer response positively affected the quality of students’ writing.

In the study *How Does Peer Editing Improve Students’ Quality of Writing*, Siew (2011) found that students who did peer editing produced better compositions from the useful feedback, a better understanding of each other’s mistakes, and a good honing of their editing skills which will have long-term effects. The teacher’s clear discussion of the editing process, modelling, and the preparation of checklists or rubrics helped make the peer editing activity a fruitful one.

Being aware of the students’ failure to understand and pay attention to teachers’ responses to students’ written work, Fahriyana (2005) conducted a survey on the value of peer editing in a writing class and found that students responded positively to peer editing and considered it very valuable. It is then suggested that an experimental study on the effect of peer editing on students’ writing be conducted. Also, it was found that the peer evaluation process resulted in improved motivation in writing and writing quality (Katstra, Tollefson & Gilbert, 1987; Philpot, 1987). Pasca (2011) investigated the effects of students’ editing done in pairs and found that it significantly improved the quality of students’ expository writing. Diab (2010) compared the effects of self-editing and peer-editing and found those subjected to peer editing significantly improved their revised drafts. Peer-editing is then recommended in the writing classroom.

Zheng (2007) studied the extent of students’ ability to correct language errors in collaborative teams as well as the teacher’s role in the activity and concluded that certain errors which are beyond the language proficiency level of students remain uncorrected. It is then suggested that some teacher’s intervention must be included to make the activity an interactive one.

Indeed, peer group evaluation of writing outputs is gaining popularity over the traditional teacher correction method and the usual self-editing method. However, such an evaluation mode is not without challenges. For one, some students do not easily agree with the comments or suggestions of their peers. Another thing, not all comments or suggestions of the peers are correct and some errors to be addressed are beyond the language proficiency level of the student-editors (Zheng, 2007). More importantly, teacher-mediation in the process counts a lot. As mentioned, peer editing, if not properly mediated by the teacher can be a waste of time. The researchers still feel the need for further investigations on the effects of varied peer editing techniques on students’ writing. This study focused on how the peer editing process can be successfully done in the writing classroom.

**Objectives**

The study aimed to determine the effects of varied evaluation techniques on the writing performance of students. Specifically, it aimed to determine the effects of the selected evaluation techniques on the students’ essays, and to identify which evaluation method results in better writing performance of students.
Methodology

Research Design
The study utilized a pre-test/post-test comparison group quasi-experimental design with the level of significance (alpha) set at 0.05.

Locale and Respondents
The study was conducted among the first year college students of Central Mindanao University, Musuan, Bukidnon, Philippines. The sample was comprised of students enrolled in English 12 (Writing in the Discipline) during the second semester of the study year of 2011-2012. For the purpose of convenience, only those writing classes handled by the lead author were considered. A total of 88 students from three intact English 12 classes participated in the study.

Instrumentation
The study utilized an expository essay writing test as the main instrument. The essay was entitled “What is a Responsible Student?” The essay should consist of five paragraphs, the first paragraph containing the thesis, the following three paragraphs as the support to the thesis, and a concluding paragraph.

A holistic rubric adapted by the teacher was used by the students in the evaluation process. A total of 88 essays were evaluated considering three features of the compositions, namely: content (focus and support), coherence (organization) and accuracy (grammar and mechanics). The components were allotted 20, 10, and 10 points respectively. The writing performance improvement was measured through the average holistic scores of students in the post-test in comparison with their scores in the pre-test.

Procedure
To achieve homogeneity in the groupings, each class was given a diagnostic writing test at the beginning of the semester to identify the students’ writing proficiency level. The students were then ranked according to their scores on the diagnostic test. Based on their ranks, students were grouped into fours and each member in every group was randomly assigned to the different evaluation groups. After the groupings, the students were introduced to the stages in the writing process, namely: the prewriting stage, the drafting stage, and the rewriting stage.

Draft writing was introduced to students as a well-structured activity. The teacher provided the topic for the students to write on and students were made to create a rough outline of the draft by letting students write five entries in their outline: a topic sentence, stating the thesis or the main point or the controlling idea of the composition; three supporting sentences to the thesis and one concluding sentence. Using the rough outline, students came up with a one-paragraph composition. For students to check their focus, they were made to underline the main point in their paragraphs.

In introducing the evaluation methods, the researcher exposed the students to four solo paragraph writing occasions after which they employed one particular evaluation method for each draft before revision. Draft writing and editing were done in the classroom. Draft revision was given as a take-home activity to give students enough time to consider feedback given and to polish their work. After such exposure, students were given a five-paragraph expository essay writing activity. The first draft as the output of this activity was the pre-test. The essay was entitled, “What is a Responsible Student.” Students were made to underline the topic sentence for each paragraph.
After the drafting, students evaluated their essays using the evaluation method assigned to them. A total of 88 essays were evaluated. Four evaluation methods were employed in the study: self-evaluation, peer (one-on-one), peer group 1 (where editors consider all aspects of the essays), and peer group 2 (where editors only consider certain aspects of the essay). All student-editors were guided by an adapted rubric/evaluation sheet prepared by the teacher.

After getting feedback, students did ratiocination on their own essays. Ratiocination is defined as a color-coding process to aid in editing and done when students are about to prepare the final paper (“Ratiocination,” 2010). With the use of highlighters or colored pens, students were made to check sentence construction, punctuations, repeated words, parts of speech, capitalization and spelling before finally revising their drafts. After the editing and the ratiocination activity, students then revised their drafts. The revised or the final draft was the post-test. Both the first draft (pre-test) and the final draft (post-test) were rated by the teacher using the same rubric used by the students in the editing sessions. For inter-rater reliability, ten essays from each group were rated by three English instructors.

**Statistical Tools**

Writing performance was measured through the average holistic scores of students. The data were summarized using means and standard deviations. A paired t-test was performed to determine whether the evaluation methods were effective. In determining which of the evaluation methods results in better writing performance, analysis of variance (ANOVA) was used. Assumptions for the analysis of variance were checked before performing the analysis of treatment effects. The Levene’s test for homogeneity of variances was performed for each ANOVA and a significant ANOVA was followed by the Tukey procedure for multiple comparisons of means.

To establish how reliable the mean of the ratings of the three raters is, an intraclass correlation (ICC) was computed. Result of a two-way mixed consistency ICC revealed that 80% of the variance of these raters is real.

**Results and Discussion**

**Effect of the Selected Evaluation Methods on the Students’ Essays**

In determining the effects of the evaluation on the students’ essays, a paired t-test was performed. Results are shown in Table 1.

<table>
<thead>
<tr>
<th>Essay Feature</th>
<th>Pre-Test M</th>
<th>SD</th>
<th>Post-test M</th>
<th>SD</th>
<th>Difference M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>9.39</td>
<td>2.89</td>
<td>10.55</td>
<td>2.55</td>
<td>1.16</td>
<td>1.73**</td>
</tr>
<tr>
<td>Coherence</td>
<td>4.27</td>
<td>1.58</td>
<td>5.11</td>
<td>1.26</td>
<td>0.84</td>
<td>1.13**</td>
</tr>
<tr>
<td>Accuracy</td>
<td>7.19</td>
<td>2.39</td>
<td>11.49</td>
<td>2.15</td>
<td>4.30</td>
<td>1.70**</td>
</tr>
<tr>
<td>Total Score</td>
<td>20.85</td>
<td>6.03</td>
<td>27.15</td>
<td>5.28</td>
<td>6.30</td>
<td>3.46**</td>
</tr>
</tbody>
</table>

Note: ** significant at 1%

The average differences between post-test and pre-test scores of students were statistically different on all of the three components of the rubric. The average increase in content (M=1.16, SD=1.73, N=88) was statistically significant, t (87) =6.292, p<.001 which suggests that the evaluation methods were effective in improving this aspect in the students’ essays. This implies that with the help of the editing methods, students were able to produce a unified composition.
with a clear focus and adequate support to their thesis. Zheng (2007) had cited that ideas and content are “higher order concerns” in essay writing. This significant result then affirms that collaborative peer editing can help students develop critical thinking skills and perform at “higher intellectual levels” (in Dooly, 2008, p.2).

Likewise, the average increase for organization (M=0.84, SD=1.13, N=88) was statistically significant, t (87) =6.960, p=.001 which also suggests the effectiveness of the evaluation methods in improving this aspect in the students’ essays. Like content and ideas, organization is among the high-order concerns in essay writing; hence, the significant result also shows the positive impact of the evaluation methods in honing higher level thinking skills of students.

Furthermore, the average increase for mechanics and grammar (M=4.30, SD=1.70, N=88) was statistically significant, t (87) =23.755, p=.001 which suggests that the evaluation methods also improve the quality of the students’ essays in terms of grammar and mechanics. This can be attributed to the fact that in the editing stage, students basically consider this aspect. Although linguistic errors belong to the low-order concerns in essay writing (as cited by Zheng, 2007), accuracy in grammar and mechanics can make the essay more readable and more engaging to the reader. Hence, the role of peer evaluation in minimizing linguistic errors should be acknowledged.

It must be noted, however, that error correction may pose some problems in the process of peer evaluation. Some students may not consider the comments of their peers and not all comments are always correct, and more importantly, some linguistic errors are beyond the language proficiency level of the students (Zheng, 2007). But as the students are given enough time to discuss their responses to each other’s outputs, such peer interaction will eventually result in cognitive development (Diab, 2010). Also, studies have shown the preference of students for this evaluation mode and that it improves students’ motivation in writing and the quality of writing (Fahriyana, 2005; Katstra, Tollefson, & Gilbert, 1987; Philpot, 1987). Hence, despite the challenges that may arise during peer editing sessions which may make the task time-consuming and seem unrewarding, the findings that the peer error correction had positively affected the technical aspect of the students’ papers should encourage writing teachers to design mediation techniques for the activity to be successful.

Finally, it was also found that the average increase in the total score (M=6.30, SD=3.46, N=88) was statistically significant, t (87) =17.078, p=.001. This result confirms the research findings that peer editing had a significant impact on students’ motivation and writing achievement which can be attributed to students’ becoming more responsible for their own learning (Fahriyana, 2005; Blasingame & Bushman, 2005; Cho, Schunn, & Wilson, 2006; Katstra, Tollefson & Gilbert, 1987; Philpot, 1987)).

This finding then implies that peer editing indeed can be an effective strategy to address the problem of giving immediate corrective feedback on students’ writing outputs which writing teachers in huge writing classes cannot possibly accomplish. This finding also implies that writing teachers should design writing instruction which will allow students enough time to evaluate their own work. Doing so, students can maximize the value of sharing and exchanging insights achieved by a fruitful interaction among them. Also, evaluating their own writing outputs can harness their critical thinking and evaluation skills (Blasingame & Bushman, 2005; Cho, Schunn, & Wilson, 2006; Felder & Brent, 2007; Nelson, 2002).

Since all the evaluation methods utilized a checklist adapted by the teacher as a guide in the editing process, and all evaluation methods have a significant effect on the essays, it can be
inferred that the teacher-mediation in the editing process contributed to such impact (Berg, 1997; Zheng, 2007). This affirms the value of using checklist or rubrics in assessment or evaluation activities (Siew, 2011).

**Evaluation Method Resulting to Better Writing Performance of Students**

In determining which evaluation method resulted in a better writing performance of students, analysis of variance (ANOVA) was used. The results are shown in Table 2.

<table>
<thead>
<tr>
<th>Essay Feature</th>
<th>Self (N=23)</th>
<th>One-on-One (N=24)</th>
<th>Peer-Group 1 (N=20)</th>
<th>Peer-Group 2 (N=21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>9.22 (2.04)a</td>
<td>9.79 (2.78)ab</td>
<td>11.25 (2.24)bc</td>
<td>12.19 (1.99)c</td>
</tr>
<tr>
<td>Coherence</td>
<td>4.65 (1.47)a</td>
<td>4.92 (1.21)ab</td>
<td>5.25 (1.07)ab</td>
<td>5.71 (1.06)b</td>
</tr>
<tr>
<td>Accuracy</td>
<td>11.00 (2.39)</td>
<td>11.25 (2.35)</td>
<td>11.35 (1.79)</td>
<td>12.43 (1.78)</td>
</tr>
<tr>
<td>Total Score</td>
<td>24.87 (5.26)a</td>
<td>25.96 (5.54)a</td>
<td>27.85 (4.58)ab</td>
<td>30.33 (3.94)b</td>
</tr>
</tbody>
</table>

Note: In a row, means followed by a common letter are not statistically different at \( p < .05 \) using the Tukey HSD procedure.

The average scores for content were different across the evaluation methods, \( F (3, 87) = 7.163, p < .001 \). The mean post-test score for peer-group 2 evaluation method was significantly higher than those of self-evaluation and one-on-one evaluation methods but not significantly higher than that of peer-group 1 evaluation method. This means that the peer-group evaluation methods have a stronger impact on the content quality of the students’ essays. This can be attributed to more insights gained by students as they engage in direct interaction with their peers during the editing process.

Such finding was also true for organization, \( F (3, 87) = 3.084, p = .032 \). The mean post-test score for peer-group 2 evaluation method was significantly higher than that of self-evaluation method but not significantly higher than that of one-on-one evaluation method and peer-group 1 evaluation method. Worthy of note is that the mean post-test scores for one-on-one, peer-group 1, and peer-group 2, which are all collaborative evaluation methods, were not found to be significantly different from each other but significantly higher than that of the self-evaluation method. This suggests that collaborative evaluation methods really work better than the non-collaborative technique which confirms the findings of Diab (2010) and Kasule and Lunga, (2010) that students who did peer editing gained better scores in their final drafts compared to those students who did editing individually. As mentioned, organization is among the higher order concerns in writing; hence, this significant result affirms the significant role of collaborative peer evaluation in harnessing students’ skills in sequencing ideas logically.

However, it could be noted that the average scores for accuracy were found to be not significantly different across evaluation methods, \( F (3, 87) = 1.920, p = 0.133 \). This finding proves to be interesting in the sense that basically editing is intended to check grammar and mechanics. In draft writing, teachers usually put much emphasis on content and organization in their writing instruction and consider grammar and mechanics in addition to content and organization in the final manuscript. Students are also aware of this matter. This finding then seems to show that giving students time to edit and rewrite their work can really help them come up with a better written product. In addition, this non-significant result can be attributed to ratioception which students did in their own essays after getting feedback. Using coloured pens and highlighters, students were able to check their grammar and other technical aspects in their own writing regardless of what editing technique they were employing.
Interestingly, the average holistic scores of students were found to be different across the evaluation methods, \( F(3, 87) = 5.143, p=.003 \). As shown in the table, the average holistic scores of peer-group 2 were significantly higher than those of the self and one-on-one evaluation method. Mean scores for the self and one-on-one evaluation methods were not found to be significantly different from each other. Likewise, mean scores for the peer-group 1 and peer-group 2 evaluation methods were not found to be significantly different from each other. This finding again shows the stronger impact of collaborative (group) evaluation methods compared to self-evaluation which is basically done solo and to one-on-one evaluation method wherein only two people exchange ideas. Hence, the expression “the more, the better” seems to be true in the editing process.

To sum up, it can be said that except for mechanics and grammar which could have been polished through the ratiocination activity as part of the methodology of the study, the peer-group evaluation methods significantly affect the writing performance of students which is in consonance with what the cited literature has found regarding the benefits of collaborative editing methods (Nelson, 2002; Hyslop, 2010; Karegianes, 1980; Siew, 2011). This finding then affirms the value of collaborative learning tasks in the writing classroom and suggests the necessity of incorporating collaborative assessment activities in writing instruction.

Conclusion

Based on the results of the study, it is concluded that the evaluation methods had a significant impact on the students’ essays. This means that letting students evaluate their own outputs enables them to see their own errors, thus harnessing their critical thinking skills. This poses a challenge for language teachers to design writing activities which can help students to become more independent and more responsible for their own learning. Among the evaluation methods, the peer group or the collaborative evaluation method resulted in better writing performance. This finding reaffirms the value of collaboration among students especially in the cognitively demanding task of writing. It was also highlighted in the study that the teacher’s mediation in the process (by providing editing checklists or by giving specific editing instructions, like the ratiocination activity) really counted a lot. This finding should encourage language teachers or writing teachers in particular, to continue exploring novel approaches in getting students more involved in the learning process. The study dealt only on editing which is one of the stages in the writing process. While student collaboration in the prewriting stage and evaluation stage may work well if properly mediated by the teacher, it would be more challenging if students are asked to come up with a collaborative writing output. Hence, other studies can be conducted on how collaboration can be successfully done in the drafting or writing stage. Such investigations will prove to be invaluable to teachers in helping students to be more prepared to function in the outside world where collaborative efforts are of paramount importance over individual contributions.

Acknowledgement
We wish to thank Central Mindanao University, Musuan, Bukidnon, Philippines for the fund to carry out this study.
References


Fahriyana, I. (2005). *Peer editing in the process of learning writing done by the fourth semester students at the English Department of Muhammadiyah University*. Retrieved from http://eprints.umm.ac.id/197/


