Peer Written Corrective Feedback on E-Mails: A Comparison of Static, Dynamic, and Integrated Partnership

Niloofar Seyed Golshan¹, Vahid Hedayati²

¹Young Researchers & Elite Club, Science and Research Branch, Islamic Azad University, Language Department, Damavand, Iran
²Shahid Beheshti University, Faculty of Language and Social Studies, Tehran, Iran

Email address:
golshanniloofar@yahoo.com (N. S. Golshan)

To cite this article:

Abstract: That learners could gain benefit from written corrective feedback provided by their peers has already been established by a large bulk of empirical research and rational argumentation. However, to what extent peer feedback is effective, and what type of peer feedback provision could best cast positive results on the writing proficiency of L2 learners is still a heat topic of debate and controversy. This study, however, looks at the quality of the peer feedback provision from a whole novel angle, and that is drawing a comparison between the peer written corrective feedback which is always provided by a fixed partner, by a varying randomly assigned partner, or by a hybrid of both fixed and random partners for the first and the second half of the treatment period. To investigate this empirically, three 35-member groups of intermediate learners were assigned to the three modes of feedback provision, and received a treatment of 22 sessions. The Analysis of the Variance on the post-test results depicted the existence of a statistically significant difference between the progress rates of L2 writing progress in the learners of the three groups. The Post-Hoc Scheffe Test determined that the significant difference was between the fixed partner and random partner groups, and also between the fixed partner and the integrated partner group. Overall, the learners who received written corrective feedback from various partners in the whole treatment period staged the most successful performance in the post-test.

Keywords: Fixed Partner, Integrated Partner, Peer Feedback, Random Partner, Written Corrective Feedback

1. Introduction

Without a single shred of doubt, the world of SLA has witnessed a metamorphosis in the quality of attention the researchers paid to the skill of writing in the course of the past 30 years, in a way that 1980s and 1990s have been decades when L2 writing research experienced some of its firsts such as the first journal allocated exclusively to L2 writing (Leki, Cumming, & Silva, 2008). Of all the arenas in which research into writing takes place, no doubt, written feedback has always been one of the most popular ones with the scholars (Evans, Hartshorn, McCollum, & Wolbersberger, 2010).

Corrective feedback simply refers to any type of feedback that a learner is furnished with in which evidence for learner error of language form exists (Russell & Spada, 2006). In field of SLA writing, there are a large number of studies in which a wide range of characteristics in learners’ L2 written texts and how those texts are responded to in terms of ideas, relevance, rhetoric organization, grammar, collocations, word choices, spelling, punctuation, and other language components (Conrad & Goldstein, 1999). The most outstanding controversy that lies in this huge bulk of research is the degree to which written corrective feedback could be effective. Some scholars, their most outstanding being Truscott (1996) argue that not only does corrective feedback on L2 writing not have any actual positive impact on learners’ accuracy, but it could also be counterproductive, to students’ writing development. On the contrary, there are scholars such as Ferris (2002) and Goldstein (2005) who have delivered iron-cast empirical evidence that judicious and purposeful written corrective feedback could make contributions to L2 writing proficiency.

As a fanatic of the latter position, the research of this paper has embarked on this study to look into the quality of written corrective feedback provided by peers from a novel angle. The fact that peer feedback could be an excellent practice to apply while teaching writing is not news to the SLA community (Mendonca & Johnson, 1994). In fact, that
learners could socially, cognitively, motivationally, and methodologically benefit from this type of feedback has already been established by research (Villamil & De Guerrero, 1996). Hyland (2000) studied the interactions of learners in a writing workshop and discovered students freely worked with each other and assisted each other’s progress by offering advice and revision notes. Hence, he concluded that the informality of peer feedback was one of its positive facets. Moreover, Rollinson (1998) and Caulk (1994) discovered that learners in their studies provided valid and proper corrective comments in their peer feedback studies. Berg (1999) and Chaudron (1984), on the other hand, yielded research findings that proposed peer feedback was a complementary element to teacher feedback since it was more focused and more specific.

To the best of the researcher’s knowledge, one angle of the peer feedback that has not been the recipient of much attention in the above-mentioned and many other peer feedback studies is the temporariness or the permanency of the peer partners. The researcher considered situation A, where learners develop a number of writing papers and they are all responded to by the written corrective feedbacks of a fixed partner in the class, and situation B, where learners develop a number of writing papers and they are responded to by the written corrective feedbacks of different randomly-selected partners in the class. The researcher also considered situation C, where some of the writing papers of learners are responded to by the written corrective feedbacks of the same partner, and some of them are responded to by varying randomly-assigned partners. Assumingly, learners in these three situations would receive feedback in different levels with different qualities, and hence they would experience different rates of progress in their L2 writing proficiency. This study, thereafter, set out to compare the impact of peer written corrective feedback. To make it possible, the researcher formulated the following research question:

**RQ:** Are there any statistically significant differences between the impact of peer written corrective feedback provided by static, dynamic, and integrated partners on learners’ L2 writing proficiency?

And to lead the experimental path of this study, the researcher formulated the research question that follows to be investigated empirically:

**RH:** There are not any statistically significant differences between the impact of peer written corrective feedback provided by static, dynamic, and integrated partners on learners’ L2 writing proficiency.

This study is significant in two major aspects: first, it would uncover which style of peer feedback provision – always provided by a fixed partner or always provided by varying partners – more efficiently impacts learners’ L2 writing proficiency. Second, since there is a hybrid treatment group in this study – in which the learners receive feedback from a fixed partner in the first half of the treatment period and from randomly-assigned partners in the second half of the treatment – it could uncover whether or not an integration of these two extreme types of feedback provision could lead to better results.

### 2. Research Design and Methodology

This study was an experimental one in its nature, which enjoyed a pre-test treatment post-test research design, and since it involved the formation of three treatment groups without the existence of a control group, it was a quasi-experimental investigation. Thus, the design of this study could be illustrated as in the following:

<table>
<thead>
<tr>
<th>Group</th>
<th>Partner Type</th>
<th>Treatment Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>Static</td>
<td>X</td>
</tr>
<tr>
<td>R2</td>
<td>Dynamic</td>
<td>X</td>
</tr>
<tr>
<td>R3</td>
<td>Integrated</td>
<td>X</td>
</tr>
</tbody>
</table>

The only dependent variable of this study was the progress of learners’ writing proficiency which was led by the only independent variable of static and dynamic peer feedback partnership. The control variables of this study were intermediate English proficiency level of the participants as well as conducting the study via e-mails in the virtual atmosphere. There were no moderator variables in this study, and intense attempt was made to reduce the role of extraneous variables that tamper with the naturalness of the data to the minimum possible level.

The participants of this study were 105 Iranian EFL learners who were selected on a random basis, and homogenized from among an initial population of 138 potential learners. All the participants were of intermediate proficiency level, adults (with the age range of 21 to 38), and they all spoke Persian as their native language. None of the participants had a career in L1 writing (e.g. journalism or book editor), and none of the participants had ever taken an exclusive L2 writing class prior to entering this study. Hence, the researcher was reassured that no member of the population was a writing professional.

In order to homogenize the participants, as well as to collect the pre- and post-treatment data, the researcher utilized one instrument only: a past version of Academic IELTS test, which was obtained from Cambridge IELTS 9 (Cambridge ESOL Local Examinations Syndicate, 2013). Since this was a past paper of the IELTS test, the researcher was sure than the testing experts in Cambridge ESOL Local Examinations Syndicate had already checked its psychometrics, and there was no need for the researcher to check its validity, internal consistency, or item functionality. It should not be left unmentioned that the researcher applied the Task 2 Band Descriptor of IELTS (Public Version) (retrieved from www.ielts.org/pdf/UOBDs_WritingT2.pdf) to score the pre- and post-test writing section of the IELTS test.

After the administration of the IELTS test, homogenizing the participants, and randomly assigning them to three experimental groups, the researcher commenced the administration of the treatment. The learners in all three groups took writing training in 22 105-minute sessions based on the course book College Writing, from Paragraph to Essay (Zemach & Rumisek, 2003).
The first session of the treatment was allocated to the briefing as to how the peer feedback should be offered – based on the writing components in the Task 2 Band Descriptor of IELTS – and what the syllabus of the course would be. The final session of the treatment was allocated to the post-test essay writing. In the course of the 20 sessions, learners were assigned 20 essays to write and email to the researcher (as the instructor of the three classes). Since the learners were not told they were participating in an academic study, this study could be considered a blind study. Moreover, the researcher used e-mail as the medium of handing in the assignments and giving the feedback back to the students, and Microsoft Word (the option track changes) was the tool the learners utilized to provide feedbacks. This matter prevented learners to notice whether the same or a different partner has provided feedback on their essays from the appearance of the feedback such as color, handwriting, underlining, and the like. Likewise, learners did not know whether the essays they are giving feedback on belonged to the same person or to different people from appearance. Hence, learners did not have any idea whose essay they were revising, and who had revised their essays. As a result, this study had a double blind design.

In the fixed partner group, all the twenty essays of each learner was revised and corrected by the same partner, and each learner corrected the essays that all belonged to the same peer learner. In the random partner group, in spite of random selection of the partners, the researcher planned the partner assignment in a way that each time a different partner is set for each learner. In other words, after one learner gave feedback on another learners’ essays, he was crossed out of the options for the next random partner selection. This way, the researcher could confidently report that each essay was revised and corrected by a different partner. In the integrated group, however, the first 10 writings of each learner were corrected by the same partner, and the second 10 essays by a random partner (with the same selection pattern as in the random partner group).

At the end of the treatment, on the 22nd session, the researcher administered the same pre-test writing test, this time as the post test. The administration of the same test was approved in terms of research principles because the gap between the two tests was longer than two weeks (Hatch & Farhady, 1981). The post-test results were, then, statistically analyzed.

3. Findings and Data Analysis

Initially, the researcher administered a past paper of the IELTS test to homogenize the participants of this sample. As Table 1 depicts, the mean was 5.94 and the standard deviation was 0.52. Hence, all the learners with scores within 5.5 and 6.5 were considered to have almost equal proficiency level in general English, and entered the study.

<table>
<thead>
<tr>
<th>Table 1. Homogenizing IELTS Test Descriptive Statistics.</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogenizing IELTS Test</td>
<td>138</td>
<td>5.0</td>
<td>8.5</td>
<td>5.949</td>
<td>.5207</td>
<td>.271</td>
<td>.914</td>
</tr>
</tbody>
</table>

The writing section of the IELTS was the pre-test of this study, as well, whose scores are illustrated in Table 2. The mean of the sample was 5.5, and the standard deviation was 0.804. Besides, Skewness was 0.239, which depicted the data was normally Skewed, and therefore parametric.

<table>
<thead>
<tr>
<th>Table 2. IELTS Writing Pre-Test Descriptive Statistics.</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing Pre-Test</td>
<td>105</td>
<td>4.0</td>
<td>7.0</td>
<td>5.05</td>
<td>.8041</td>
<td>.647</td>
<td>.239</td>
</tr>
</tbody>
</table>

Table 3 depicts the mean of the randomly assigned three treatment groups, which are 5.68, 5.4, and 5.44. It goes without saying that the average writing scores of the participants in the three groups are almost equal.

<table>
<thead>
<tr>
<th>Table 3. Three Treatment Groups’ Pre-Test Descriptive Statistics.</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Partner Group</td>
<td>35</td>
<td>4.0</td>
<td>7.0</td>
<td>5.681</td>
<td>.9498</td>
<td>.902</td>
<td>.108</td>
</tr>
<tr>
<td>Random Partner Group</td>
<td>35</td>
<td>4.5</td>
<td>6.5</td>
<td>5.400</td>
<td>.6039</td>
<td>.365</td>
<td>.089</td>
</tr>
<tr>
<td>Integrated Partner Group</td>
<td>35</td>
<td>4.0</td>
<td>7.0</td>
<td>5.443</td>
<td>.8023</td>
<td>.644</td>
<td>.016</td>
</tr>
</tbody>
</table>

Moreover, in order to make sure that the three groups had learners with equal writing proficiency, the researcher used the Analysis of the Variance. As Table 4 demonstrates, the p value was 0.319 (p < 0.05 = significant), and hence no statistically meaningful difference was reported between the three groups, and their participants enjoyed almost equal writing skill. This meant that the sample was homogenized and equalized, and ready for the administration of the treatment.
After the administration of the post test, the researcher calculated the descriptive statistics, and as Table % signifies, the Skewness was 0.199, which demonstrated the data was normally Skewed, and hence parametric. The mean was 6.09, which showed 0.585 improvement compared to the mean of the pre-test. This indicated that all the learners in the sample improved in their English writing skill as a result of the peer feedback provision.

Taking a glance at the descriptive statistics of the post-test scores in Table 6, it is noticed that the smallest mean belonged to the fixed partner group (5.77), while the learners in the random partner group managed to earn the largest mean (6.34). The 6.15 Integrated group learners’ mean stood in the middle, as well. The standard deviation of the random partner group was the largest (6.34), which signified that the learners in this group had the highest degree of deviation from the mean.

The Analysis of the Variance between the post-test scores of the sample, as illustrated in Table 7, depicted the \( p \) values of 0.000 (\( p < 0.05 = \text{significant} \)) which reported a statistically meaningful differences between the results of the three groups. Thereafter, the research hypothesis of this study is rejected, and the data analysis vividly indicated there is a statistically significant difference between the impact of peer written corrective feedback provided by static, dynamic, and integrated partners on learners’ L2 writing proficiency.

So as to locate the significant difference that the Analysis of the Variance depicted, the researcher administered the Post-Hoc Scheffe Test on the post test scores. As Table 8 illustrates, the \( p \) value between the fixed partner group and the random partner group was 0.000 (\( p < 0.05 = \text{significant} \)), and the Mean Difference was -0.571 (Mean Difference \(< 0.05 = \text{significant} \)). Hence, there was a meaningful difference between the learners of these two groups. Besides, since the \( p \) value between fixed partner group and integrated partner group was 0.016 (\( p < 0.05 = \text{significant} \)), and the Mean Difference was -0.385 (Mean Difference \(< 0.05 = \text{significant} \)). Hence there was a significant difference reported between these two groups, as well. However, with the \( p \) value of 0.373, and the Mean Difference of 0.185, no meaningful difference was found in the post-test performance of random partner and integrated-partner groups.

The analysis of the means for groups on homogeneous subsets on Scheffe test, on the other hand, revealed that the
integrated partner and the random partner groups had almost equal homogeneous subsets, while the fixed partner group was pinpointed in another subset. This depicted fixed partner group had a significant difference on statistical mean with them.

**Table 9. Means for Groups in Homogeneous Subsets on Scheffe Test.**

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Subset for alpha = 0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Fixed-Partner Group</td>
<td>35</td>
<td>5.771</td>
</tr>
<tr>
<td>Integrated-Partner Group</td>
<td>35</td>
<td>6.157</td>
</tr>
<tr>
<td>Random-Partner Group</td>
<td>35</td>
<td>6.343</td>
</tr>
<tr>
<td>Sig.</td>
<td>1.000</td>
<td>.373</td>
</tr>
</tbody>
</table>

Overall, it could be argued that learners in the random partner group had the largest extent of progress in their L2 writing proficiency, and those of the fixed partner group had the lowest progress rate.

4. Discussion and Conclusion

As it was discussed earlier, the statistical analysis of the post-test results of this study revealed that learners who received feedbacks from varying randomly selected partners staged a more successful writing performance compared to the ones who received their feedback all from the same fixed partner or even the ones who enjoyed a hybrid position of receiving feedback from both fixed and varying partners for the first and second half of the treatment period. Several different reasons could be considered for this significant difference.

First, as Lockhart and Ng (1995) proposed, the type of feedback the students receive on their writing papers determines the approach they tend to take toward writing. Hence, when learners are exposed to various styles of correction, they figure out diverse approaches that lie ahead of them to writing, and this variety enriches their writing power.

Second, Rollinson (2005) proposed that Peer feedback, in its very nature, was a type of interaction between the reader and the writer, and through peer feedback, a cooperative conversation took place between these two parties via which the meaning was negotiated. In the course of this study, the learners who received feedback from fixed partners were engaged in negotiation dialogs with one person throughout the whole period, and they merely had one person to assist them become better writers. It goes without saying that a mutual dialog with one fixed person on a fixed subject has an expiration data with is not very distant from the beginning comment, and after a short while, the two speakers run out of ideas. Meanwhile, those in the random partner feedback had a chance to make conversation with 20 different people, and each person opened up the dialog from a different angle. So it is natural that the latter group gain more mastery in the course of the treatment. Barnes (1976) added that this increased chance of interaction through feedback with multiple learners opened up new socio-cognitive points of view.

In the same regard, Hyland and Hyland (2006) stated that peer feedback engages learners in an "exploratory talk" in a "formative developmental process" (p. 6) which improves learners’ writing proficiency through the interpretation and revision of other learners’ writing. Naturally, a wider variety of essays from a wider variety of writers does one go through and comment on, a broader scope of material s/he is exposed to, and the greater amount of knowledge s/he would earn from them. As Mittan (1989) added, the more the learners promote their analytical skills, the better producers they become.

In a totally different view, the greater progress of the random partner group could be justified by Grabe and Kaplan’s (1996) comments as to responding to peers’ work engages learners in each other’s writing, in a way that they could observe similar shortcomings and weaknesses in the essays of their own. They could see what they have done better in their own writing than the writer of the essay there are going through, and compared to this writer, what areas they need to improve (Witbeck, 1976). It could, thus, be concluded that in the fixed partner group, the learners only had one person’s writings to see and reflect on, but those in the random partner group had the luxury of going through 20 writings from 20 different people, and this way, they managed to reflect on them and compare them to their own writing style from 20 different angles.

One more reason to justify the findings of this study is that, at times, learners might not welcome criticism from their peers very warmly, and may respond defensively to their peers’ feedback (Amores, 1997). The probability of this happening in a fixed partner setting is far more than in the random partner setting. If learners believe the fixed partner who always criticizes their writing is less knowledgeable than them, is picking on some areas of writing and neglecting others, is biased, or aims at criticizing them to prove his/her own superiority rather than help them flourish and thrive, they might fall cross with their partners, and do not care much about the feedback. However, in the random partner group, this might happen with the feedbacks by one or maximum two partners, but not throughout the whole period. In this study, the learners were double blind, and in case of effectiveness of this reason, it would have certainly had unconscious impact on the learners.

An interesting facet of the findings of this study is that they draw a rejection line on the comments of Carson and Nelson (1994) and Nelson and Murphy, (1992) who had argued that there was weak interaction between the peer groups in an L2 class due to the different cultural and educational backgrounds. Hence, the poverty of interaction between the peers would weaken the impact of peer feedback on L2. If these comments were true, the fixed partner group would have gained better results because after a while of commenting on each other’s writings, the two peers would unconsciously get to know each other and would unconsciously learn to live with their differences. So they would have gained much better results compared to those in the random partner group who each time faced comments from a totally different person and an essay from a totally
different person to comment on.

5. Limitations of This Study and Recommendations for Further Research

A very good idea for a follow-up study to this investigation is repeating the same design with self-conscious participants. Ferris (2003) proposed that based on research, some learners doubt the value of the feedback provided by their peers, and in some occasions, they prefer to go with the structures they had already used in their first draft rather than the going with the corrective feedback their peers had provided. In this study, this matter could not have been considered since this study was a double blind study. The participants of this study did not any idea they were part of an academic research, and they did not have any idea whether a fixed person commented on their essays of it was a different person each time. As long as the same design is repeated, and the participants are self-conscious as to whether they receive feedbacks from the same peer or a randomly assigned peer, it could be determined which group value the ideas of their peers more.

Another applicable idea for further investigation is to repeat the same design in a multi-cultural setting. Allaei and Connor (1990) argued that in multi-cultural collaborative circumstances, peer response might result in conflict or discomfort. All the participants of this study were Iranian and spoke Persian as their native language. So this matter could not be taken into account in its design. Provided this study is repeated in a multicultural or cosmopolitan atmosphere, it could be determined which style of peer feedback provision fits better in a multicultural setting.

Graham and Perin (2007) remarked that in spite of all the research that has taken place in the field of writing, still schools are not doing a good job in teaching learners how to write. This gives rise to the necessity of research to find proper ways of teaching writing to young learners. This study was conducted on adult learners; hence, no clear-cut judgment could be passed as to how properly peer feedback by randomly assigned partners could work on teenage learners. Further study could determine the functionality of this system on young learners.

Moreover, the skill of speaking is another area of language learning in which feedback from peers could provide benefits to learners. A similar study could be designed and implemented to determine whether or not the results of this study are true of the peer feedback on speaking, as well.

References


