

# Improving the Effectiveness of the Peer Feedback Technique: The Impact of Focusing EFL Student-writers on Macro Level Features

Dr. Suliman Mohammed Alnasser

Department of English Language & Literature, Faculty of Arts

King Saud University, Saudi Arabia, Riyadh

P.O. Box 361228, Riyadh 11313, Saudi Arabia

Tel: +966503448060 E-mail: Smalnasser@ksu.edu.sa

Dr. Hesham Suleiman Alyousef (Corresponding author)

Department of English Language & Literature, Faculty of Arts

King Saud University, Saudi Arabia, Riyadh

P.O. Box 50574, Riyadh 11533, Saudi Arabia.

Tel: +966553000412 E-mail: hesham@ksu.edu.sa

Received: November 23, 2014 Accepted: December 6, 2014 Published: December 10, 2014

doi:10.5296/ijele.v3i1.6764 URL: <http://dx.doi.org/10.5296/ijele.v3i1.6764>

## Abstract

The study on which this article is based suggests a new form of the peer feedback technique that focuses learners on providing their fellow students with macro peer feedback alone, leaving the provision of micro feedback to the teacher. It investigated the impact of the new form on learners' overall writing quality, and on macro and micro level writing features. The participants were 41 Saudi EFL undergraduate students undertaking an English programme. The design of the study consisted of two phases: during phase 1 the conventional form was introduced, and during phase 2 the new form was introduced to the participants. The participants developed and practised giving and receiving peer feedback on 10 argumentative essays in total, two drafts each. Pre-, mid- and post-tests and mid- and post-interviews were administered. The findings of the study suggest that each form of the technique had a significant impact on the participants at a statistical level, with the new form having even

greater impact. It was also found that the new form had a more positive impact at the macro level than the conventional form.

**Keywords:** Peer feedback technique; macro level; micro level; collaborative writing; writing features

## 1. Introduction and Background

Collaborative learning has been given increasing attention in recent years (e.g. Liu & Carless, 2006; Strijbos, Narciss, & Dünnebier, 2010; Van der Pol, Van den Berg, Admiraal, & Simons, 2008; Zheng, 2012). Its wide acceptance in higher education is a result of the need for a shift from teacher-centred to student-centred classrooms, since a student-centred environment is now considered essential for establishing active learning situations that can help to enhance learning outcomes (Orsmond, Merry, & Reiling, 2002).

Collaborative learning means that the tasks are done together whereas cooperative learning means that the tasks are divided between learners (Alyousef & Picard, 2011). Collaborative writing is widely recognised as a central component of classroom learning by researchers in the field of L2 writing (e.g. Hu & Lam, 2010; Lundstrom & Baker, 2009). Peer feedback (hereafter PF), a technique whereby students provide comments to their colleagues regarding the quality of their work, is generally categorised as a collaborative learning technique, and is frequently employed in L2 writing classrooms. From a socio-cognitive perspective, it can be seen as "...a formative developmental process that gives writers the opportunities to discuss their texts and discover others' interpretations of them" (Hyland & Hyland, 2006, p. 6).

Although the PF technique has proven to be beneficial in L2 writing classrooms, several studies have reported concerns about it. Therefore, few researchers have attempted to improve its effectiveness by introducing new forms of this technique. For example, Cho, Schunn and Wilson (2006) compared the quality of PF given by a single rater (i.e., a peer) and by multiple raters, and Gielen et al. (2010) examined two forms of PF: a question form (in which learners highlight their desire for specific comments on their written work) and a reply form (in which learners receive PF and then reflect upon it by highlighting the main factors that affected their learning). Some of these studies have contributed to enriching the quality of PF and writing performance; however, none of them involved focusing learners' attention completely on macro level issues when providing PF on writing. Although most of the studies on PF have focused on macro and micro issues at the same time, some have paid more attention to macro than to micro issues (Al-Hazmi & Scholfield, 2007; Min, 2006). To the best of our knowledge, only Van Steendam and colleagues (2010; 2014) examined an aspect of the area involving macro writing features and PF. However, their study examined the effect of different types of *instruction* (i.e., observation and practising) in revision strategies, followed by the students imitating these instructions both in pairs and individually, on the *quality* of macro feedback students provide during PF sessions. In other words, Van Steendam et al did not measure the effect of focusing on macro level features on linguistic progress. Therefore, it was the purpose of the study discussed here to introduce a new form of

PF technique with the purpose of overcoming its weaknesses and presenting it in an optimal way. The focus of this article is on the impact of this form on the overall quality of writing and on writing features. It is hoped that the results of the study presented here will contribute to research into the field of using PF technique in L2 writing.

### *1.1 What criticisms are there concerning the current use of PF?*

An examination of the literature leads to the conclusion that there are several drawbacks to the way in which conventional PF (i.e., providing PF based on macro and micro level features) is presented in L2 writing classrooms that may be jeopardising its effectiveness. For example, several studies (e.g. Broekkamp & Van den Bergh, 1996; Chenoweth & Hayes, 2001; Truscott, 1999; Van Steendam et al., 2010; Van Steendam et al., 2014) have reported a tendency on the part of learners during PF sessions to pay more attention to micro level — i.e. surface level features such as grammar and proper use of vocabulary — than to macro level features — i.e. meaning level features such as idea development and the organisation of a written text. The result of this tendency may be that learners are not equipped with the ability to write effectively at the macro level, in that they may express their thoughts poorly, even if they possess adequate linguistic skills; consequently, their learning of the knowledge they need for advanced writing courses may be hindered. Furthermore, Zheng (2012, p. 115) concluded that in PF sessions learners can act as *passive* participants (in that they tend to take a “...peripheral role, whose discourse is far from reaching a solution to the problem, assuming less important responsibility”), and thus they may not benefit properly from this technique. Moreover, some studies suggest that learners do not accept most of the feedback offered by their peers, and Min (2006, p. 133) found that an average of only 39% of such feedback was accepted at both macro and micro levels, arguing that this is the result of learners producing poor quality PF. In addition, a recent study conducted by Zhao (2010) reports that learners were more likely to accept feedback provided by their teacher than by their peers, which is not surprising, since most of the existing studies on PF in writing in ESL and EFL contexts have reported that teacher feedback was more frequently incorporated in revised drafts than PF. Various reasons for this acceptance of teacher feedback rather than PF were provided: for instance, students consider teacher feedback to be more trustworthy and professional (Yang, Badger, & Yu, 2006), and to be more significant and more useful (Tsui & Ng, 2000). This could be one of the reasons that have led the researchers to conduct a series of studies in which they have attempted to increase learners’ acceptance of PF by training learners in how to provide PF (e.g. Min, 2005, 2006, 2008). It may be surmised that these weak points of the conventional PF may be causing it to fail to be very effective in writing classrooms, and overcoming these weaknesses could increase our understanding of and give us deeper insights into the entire PF process.

### *1.2 What is new about PF in this study?*

In recent years, researchers have made several attempts to enhance the use of the technique and overcome its weaknesses in order to produce better outcomes in the learning process. The study discussed here, for instance, attempted to learn how the practice of PF can be used in an optimal way in a Saudi context in order to make the learning process more efficient. In other

words, the intention was to improve learners' overall writing quality and to improve their understanding of what writing involves (i.e., raising writing awareness). It is hoped that the carefully designed techniques for using PF introduced in the study will enrich the literature and add genuine value to the PF process. The new suggested form of PF differs from the conventional/traditional PF in that it forces learners to give feedback based on the macro level features of writing and reserves the provision of micro level feedback for the teacher alone.

The design of the new form was inspired by the fact that learners are likely to become preoccupied with micro level features when they provide feedback to one another, the principal result of which is the neglecting of more important writing features, i.e., the macro level features (Van Steendam et al., 2010; Van Steendam et al., 2014). In addition, Ferris (2004) concludes that students are likely to be less capable of self-editing lexical errors, and thus teachers may need to employ a variety of treatments (e.g. including other sources of feedback); however, several studies have shown that learners have a preference for receiving micro feedback on their written work (Ferris, 2004; Leki, 1991). It was therefore decided that in the intervention for this research, the students would be forbidden from providing micro PF to their fellow students and that teacher micro feedback would be included as an integral part of the new form. It was thought that if the participants received micro feedback from a reliable source (i.e., the teacher), they would have more exposure to and thus acquire more knowledge of macro features. This may also allow the learners to have more confidence in macro PF and, therefore, make use of a greater proportion of PF. In general, it can be said that the workload when using the new form is shared by the teacher and his/her learners.

## 2. Methodology

The study under discussion here was a research project in which the PF technique was introduced in two different forms (the conventional form and a new form) to learners in a higher educational context. A quasi-action research approach was employed in the study using a one-group design for data collection (it should be pointed out here that it was not possible to employ a control group for many reasons, one of which was that there were no other writing sections running when the data were collected and the participants refused to be divided into two groups). The participants were 41 male EFL Saudi undergraduate learners undertaking an English language programme. The participants were in their final year and shared a similar educational background. All the participants agreed to participate in the study and signed a consent form. The researcher/teacher took over an advanced writing module and collected the data during the whole academic semester— i.e., data were collected for a period of 15 weeks. There were 3 one-hour sessions a week, which were fixed and could not be altered. Our study aimed to answer the following research questions:

*RQ1: What is the effect of each introduced form of the PF technique on students' writing?*

*Sub-RQ 1.1: What is the effect on their overall writing quality?*

*Sub-RQ 1.2: What is the effect on the macro and micro level writing features?*

A mixed method approach was used in the research, combining both quantitative and qualitative methods. Three stages of writing tests were administered in order to measure the effect of introducing each of the two treatments on the participants' writing. Additionally, three stages of identical questionnaires and two stages of identical semi-structured interviews were administered in order to investigate their views concerning the effect on writing, and to measure any changes in these views before and after practice with both kinds of the PF technique.

Specifically, the three writing tests were introduced before and after the students' exposure to each treatment. The tests took the form of writing argumentative essays of about 250 words in length. The given topics were as follows: pre-test: *Is cheating out of control?* mid-test: *Is competition good?* and post-test: *Is homework harmful or helpful?*

Moreover, two assessors were used to evaluate the pre-, mid- and post-tests using Paulus's (1999) rubrics for evaluating essays - see Lundstrom and Baker's (2009) appendices for the rubrics- and the average of the assessors' two scores for each test and student was calculated. The rubrics were carefully designed to ensure that they provided clear and detailed explanations of both macro and micro level writing features. The scoring rubrics give a scale from 1 to 10, where ten is the highest mark. The writing features were categorised on six levels: organisation, development, cohesion/coherence, structure, vocabulary and mechanics. The first three categories were grouped under macro level features, and the latter three as micro level features.

With regard to the questionnaires, a five-point Likert scale was employed, incorporating the commonly used categories for rating scales: (1) strongly disagree, (2) disagree, (3) not sure, (4) agree and (5) strongly agree. The questionnaire items were:

- 1- *Reading my classmates' essays helps me to improve the organisation of my essay.*
- 2- *My essays improve after revisions.*
- 3- *My classmates' comments in peer response sessions make the organisation of my composition worse.*
- 4- *After each revision, the organisation of my essay becomes better.*
- 5- *My writing quality will deteriorate because of peer feedback.*

With regard to the qualitative aspect of the study, 11 participants volunteered to be interviewed and were asked the following question:

- 1- *Do you think your own writing changed as a result of PF? If so, how do you think it changed, and what was it about PF that led to those changes?*

## 2.1 Practical procedures

The design of the study consisted of two consecutive phases, as described below:

### 2.1.1 Phase 1

In this phase the participants were given a pre-questionnaire and a pre-test, and then they were trained in how to use the PF technique employing Min's (2006) module, in which the teacher demonstrates aloud how to provide feedback. Next, the participants wrote an argumentative essay and generated the first drafts. After that they were assigned to work with peers who possessed the same level of language proficiency- this was determined by the researcher by referring back to their academic records. They exchanged essays and started taking turns in proving PF to each other using the evaluation form, which was adapted from Al-Hazmi and Schofield (2007) and which provides detailed questioning on macro and micro writing features, and corresponds to Paulus's (1999) scoring rubrics. After exchanging comments, the participants developed their second drafts. They wrote four additional argumentative essays, exchanging comments and developing two drafts on each topic. Therefore, the participants practised providing PF on five argumentative essays, two drafts each. After practising the conventional PF technique five times, the mid-test, the mid-questionnaire (consisting of the same questions as the pre-questionnaires) and mid-interviews were administered.

### 2.1.2 Phase 2

After administering the mid-test, the mid-questionnaire and mid-interviews, the researcher explained the difference between the conventional form and the new form of the technique. The difference was that participants would provide only macro comments to their fellow students- using the section related to macro features in Al-Hazmi and Schofield (2007) evaluation form- and the researcher/teacher would deal with the micro comments. After the difference had been clarified, the participants started developing the first drafts of their sixth essay, which were collected by the teacher and returned in the following session along with teacher micro comments. After that the participants exchanged essays and started providing only macro PF in pairs. After practising the new form on five essays (two drafts each), the post-test, identical post-questionnaires and post-interviews were administered.

In brief, the participants practised the new form for the same length of time as they practised the conventional use of the PF technique, and the change in their writing skills and views were investigated both before and after exposure to each type of this technique.

## 3. Results and Analysis

SPSS software was used to examine the quantitative data obtained from the pre-, mid- and post-tests and the pre-, mid- and post-questionnaires. This test aimed to establish whether or not there were statistically significant improvements after each of the two interventions (see Appendix A, Table 1). It should be noted here that negative questions in the questionnaire were reversed, i.e., recoded as positive questions. This was done to allow the measurement of the students' responses to all questions in the same direction (i.e., measurement in a positive direction). In order to analyse the data obtained for this study, it was decided to calculate the statistics in the form of means (*M*) and standard deviations (*SD*) and to use parametric and non-parametric tests.

The Kolmogorov-Smirnov test of normality was run on the writing tests scores, and the

results suggested that the data were normally distributed. Therefore, the parametric Paired Sample T-Test was used. The test of normality was also run on the questionnaire data and the results suggested that the data were not normally distributed. Therefore, the Wilcoxon Signed Ranks Test was used for the analysis (Gray & Kinnear, 2012; Larson-Hall, 2010).

### 3.1 Writing tests

We present in this section the findings of our analysis of the overall scores of the writing tests and the effects on macro and micro writing features.

#### 3.1.1 Overall scores

The analysis of the writing tests results (i.e., pre-, mid- and post-tests) revealed that no student scored below 4 in the post-test, while in the mid-test no student scored below 2.7. In the pre-test the minimum score was 1.6, as shown in the Table below.

Table 1. Paired Samples Statistics (Linguistic performance).

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-test	3.4309	41	1.15363	.18017
	Mid-test	4.5285	41	.98665	.15409
Pair 2	Mid-test	4.5285	41	.98665	.15409
	Post-test	6.1484	41	1.21771	.19017
Pair 3	DIFF1	-1.0976-	41	.62581	.09773
	DIFF2	-1.6199-	41	.85087	.13288

Generally speaking, a large proportion of the students scored as follows in the three tests: pre-test; between 2.4 and 3.6; mid-test: between 3.7 and 5.2, and post-test: between 5.2 and 6.6.

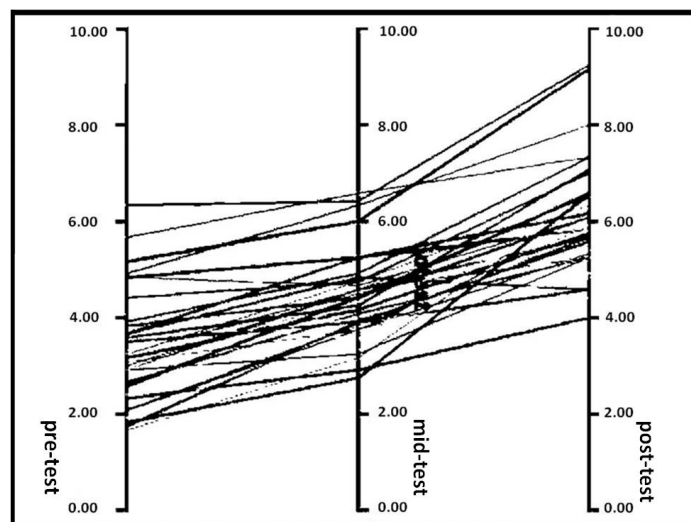


Figure 1. Students' writing tests scores in the pre-, mid- and post-tests

In order to analyse the changes in the students' writing scores after exposure to each treatment, the Paired Sample T-Test was used to determine whether or not there were differences in the students' performances in the pre-, mid- and post-tests. The results showed statistically significant improvements ( $t = -11.23$ ,  $p < 0.001$ ) after each of the two interventions (see Appendix A, Table 1). A statistically significant difference was found between the pre- and mid-test results, with a mean score of  $M = 3.43$  and a standard deviation of  $SD = 1.15$  obtained for the pre-test, and a mean score of  $M = 4.52$  and a standard deviation of  $SD = 0.98$  obtained for the mid-test. A statistically significant difference was found ( $t = -12.19$ ,  $p < 0.001$ ) between the mid- and post-test results, with a mean score of  $M = 6.14$  and a standard deviation of  $SD = 1.21$  being obtained for the post-test.

The differences found between the means obtained for the pre- and mid-test results and the differences found between the mid- and post-test results were both equally statistically significant, at  $p < 0.001$ . For this reason, it was considered to be worth undertaking a further comparison between the differences in the pre- and mid-test and the differences in the mid- and post-test results to show if the statistical significance of the change after exposure to the second treatment was greater than that of the change after exposure to the first treatment. This was calculated by subtracting the results each student obtained in the pre-test from his results in the mid-test, and then subtracting the results each student obtained in the mid-test from his results in the post-test (i.e., pre-test – mid-test = *diff1*; mid-test – post-test = *diff2*). After that, these differences (*diff1* & *diff2*) for all the students could be compared using the Paired Sample T-Test.

Therefore, a further comparison was made between the differences in the means obtained for the pre- and mid-tests (coded as *diff1*), and the differences in the means obtained for the mid- and post-tests (coded as *diff2*). The results revealed a statistically significant difference ( $t = 2.78$ ,  $p = 0.008$ ) between the first and the second differences, with a mean for the first difference (*diff1*) of  $M = -1.09$ , and a mean for the second difference (*diff2*) of  $M = -1.61$  (see Table 1 above and Appendix A, Table 1).

To conclude, the results showed that the students' writing scores had improved at a statistically significant level ( $p = 0.001$ ) after exposure to each of the two treatments (the conventional form of PF and the new form of PF). In addition, the improvement at the end of the second treatment was statistically significantly greater than that at the end of the first treatment.

Several patterns were observed in the scores of all the students in all the tests. It was observed that students who started at high and mid-levels ( $N = 8$ ) in the pre-test were likely to show greater progress in the post-test (i.e., after exposure to the new form of PF) than in the mid-test (i.e., after exposure to the conventional PF). A few students who started at mid- and high-levels made more or similar progress in the mid-test compared with the post-test. Additionally, it was found that those who started at low levels ( $N = 14$ ) were likely to make less progress in the post-test than in the mid-test. Finally, a few students who got low scores ( $N = 17$ ) in the pre-test showed more evidence of progress between the mid- and the post-test than between the pre- and the mid-test. Therefore, it can be said that there were no clear



trends among those who started at low levels.

### 3.1.2 Effects on writing features (macro and micro)

As this study used Paulus's (1999) scoring rubrics, which categorise writing features into six categories (i.e., organisation/unity, development, cohesion/coherence, structure, vocabulary and mechanics), the grades for each of these categories (on a score from 1 to 10) were used to provide a more detailed picture of the results. In other words, when the essays were marked, each essay was given both sub-scores for each of the aforementioned categories and an overall score. The sub-scores for each of these six categories were entered into the SPSS program and a Paired Sample T-Test was carried out. The results revealed a statistically significant difference in each of the six categories in the mid- and post-tests at the  $p < 0.001$  level (see Appendix A, Tables 2, 3). In other words, the scores for each sub-test were significantly higher in the mid-test than in the pre-test, and significantly higher in the post-test than in the mid-test.

For the purposes of this research, the first three categories (organisation/unity, development and cohesion/coherence) were identified as macro level features, and the remaining three categories (structure, vocabulary and mechanics) as micro level features. In order to carry out a more detailed analysis of the contribution of these categories to the development of the students' writing, therefore, comparisons were made between macro and micro level features and of how they changed in the three tests (i.e., pre-, mid- and post-tests).

The Paired Sample T-Test was used to determine whether there was a statistically significant difference between the macro and micro level features of the students' performances in the pre-, mid- and post-tests. Using SPSS, the test was run on the pre-micro features ( $M = 3.45$ ,  $SD = 1.22$ ) and the pre-macro features ( $M = 3.40$ ,  $SD = 1.16$ ); the mid-micro features ( $M = 4.43$ ,  $SD = 0.98$ ) and the mid-macro features ( $M = 4.62$ ,  $SD = 1.03$ ), and then the post-micro features ( $M = 5.87$ ,  $SD = 1.30$ ) and the post-macro features ( $M = 6.41$ ,  $SD = 1.18$ ) for all the students in the study, to establish whether there were statistically significant improvements before and after each of the two interventions (see Appendix A, Tables 4, 5). No statistically significant difference was found ( $t = 0.618$ ,  $p > 0.05$ ) between the pre-micro and pre-macro levels. However, a positive statistically significant difference was found ( $t = -2.85$ ,  $p = 0.007$ ) between the mid-micro and mid-macro levels. Additionally, there was also a positive statistically significant difference ( $t = -7.04$ ,  $p < 0.001$ ) between the post-micro and post-macro levels.

An additional analysis was carried out in order to probe deeply into the participants' results by investigating the development (difference in the change) of each sub-category after each treatment. The investigation measured whether the difference between the means of each sub-score of the pre- and mid-tests (coded as *diff1*), and that of the mid- and post-tests (coded as *diff2*), was statistically significant. The results revealed a statistically significant difference at the  $p < 0.05$  level between the first and the second differences in development, cohesion and structure, but not in organisation, vocabulary and mechanics (see Appendix A, Tables 2, 3). Therefore, this analysis suggests that the majority, but not all, of the macro features were affected positively at a statistically significant level.

In conclusion, the results showed that the students' writing improved in all the six categories described by Paulus (ibid), and to a statistically significant degree after exposure to each of the two treatments. Furthermore, a greater improvement was found in most of the macro level features than in the micro level features after exposure to each of the two treatments. In more detail, the improvement in the development, cohesion and structure categories at the end of the second treatment was greater, at a statistically significant level, than that at the end of the first treatment. In other words, the results indicated greater improvement in the majority of the macro level features at the end of the second treatment than at the end of the first treatment.

### 3.2 The questionnaire

Five items in the questionnaires enquired about the learners' views on the effect of PF on their writing. Since the interest of the researcher lay in the theme they reflected rather than in their responses to single items, these were combined as a single variable. These items were:

- 1- *Reading my classmates' essays helps me to improve the organisation of my composition.*
- 2- *My essays improve after revisions.*
- 3- *My classmates' comments in peer response sessions make the organisation of my composition worse. (reverse coded)*
- 4- *After each revision, the organisation of my essay becomes better.*
- 5- *My writing quality will deteriorate because of peer feedback. (reverse coded)*

Table 2. Effects on writing

	N	Mean	Std. Deviation	Minimum	Maximum
Pre_PF effect on writing	41	3.9561	.65040	2.40	5.00
Mid_PF effect on writing	41	4.1073	.63576	2.00	5.00
Post_PF effect on writing	41	4.1707	.60755	2.60	5.00

The Wilcoxon Signed Ranks Test was used to determine whether or not there was a statistically significant difference in the students' beliefs with regard to the effects of PF on writing in the pre-, mid- and post-questionnaires (see **Table 2** above, and Appendix A, Tables 6, 7). Using SPSS, the test was run first on the pre-questionnaire results ( $M= 3.95$ ,  $SD= 0.65$ ; with 2 out of 5 items being highly positive, and 3 items being positive) and the mid-questionnaire results ( $M= 4.10$ ,  $SD= 0.63$ ; with 3 out of 5 items being highly positive, and 2 items being positive), and then on the mid-questionnaire and post-questionnaire ( $M= 4.17$ ,  $SD= 0.60$ ; with 4 out of 5 items being highly positive, and 1 item being positive) results for all the students in the study, to establish whether or not there were statistically significant changes in their beliefs with regard to the effects of PF on their writing after each of the two

interventions. No statistically significant difference was found ( $Z = -1.68$ ,  $p > 0.05$ ) between the pre- and mid-questionnaire results. Moreover, no statistically significant difference was found ( $Z = -0.76$ ,  $p > 0.05$ ) between the mid- and post-questionnaire results. However, when the pre- and post-questionnaire results were compared, a statistically significant difference was found at the  $p = 0.01$  level and  $Z = -2.44$ .

### 3.3 *The interviews*

In the mid-interviews, the students reported seeing an improvement in their writing skills. They reported that their writing had been affected in ways that made them revise their work before submission; learn new vocabulary; discover new things, such as ways of using cohesive ties and adding more clarification to the essay; begin to pay more attention to macro level features; know how to include a thesis statement; think differently; widen their mental horizon; have more practice; correct more grammar and vocabulary; produce smoother writing; improve the style of their writing, and improve the way they delivered their message to the reader.

However, in the post-interviews more interesting findings emerged. All the students reported more and deeper improvements in their writing. None of them reported the opposite. According to them, the improvement affected their writing in various ways: for instance, learning more than one way of writing; improving the content; improving the style; conducting a revision before submission; using different words; learning different vocabulary; using more complex sentences; having better coherence; paying more attention to attracting the reader; better organisation in the introduction, and improving the presentation of the essay. It was also reported that the improvement in their writing had been transferred to other courses, the result of which was that they gained higher marks.

In brief, PF improved the students' writing. However, the improvement found after exposure to the new form of PF was greater, leading to a better effect in terms of macro features.

#### 3.3.1 *Reasons that led to these changes*

According to the mid-interviews, there were probably five main reasons for these changes (improvements) in their writing. The most commonly reported reason was the fact that they received feedback on their essays. The second reason was reading colleagues' essays and commenting on them. The third reason was having the ability to discuss and negotiate the received PF, thus encouraging two-way feedback. Fourth, there was no formality between peers, which led to their feeling comfortable with no fear, compared with the fear found to be associated with teacher-student feedback situations. Finally, using the evaluation checklist was also reported to affect their writing positively.

In the post-interviews, other factors including the aforementioned were reported as having an effect on their writing. These were: changing colleagues; receiving comments from the teacher; honesty on the part of the audience; no grades to give and "nothing to gain or lose" (indicating the formative use of the technique), and imitating proficient peers.

#### 4. Discussion

The study referred to in this paper investigated an attempt to improve the PF technique. It measured the linguistic impact of a new form of the technique on EFL student-writers and compared this with the impact of the conventional form. Starting with the administered tests (i.e., pre-, mid- and post-tests), it was found that the average score in the pre-test (i.e., before the students were exposed to either form of the PF technique) was 3.43 out of 10. This can be considered to be low relative to the highest point on the used scale, and may be indicative of the students' low achievement in their past learning. However, after they had been exposed to the conventional form of PF, a statistically significant improvement was found in the mid-test, with a mean of 4.52 out of 10. Moreover, a statistically more significant improvement was found after exposure to the new form of PF, with a mean of 6.14 out of 10, which suggests that the new form appears to have a greater positive effect on learners' writing.

These results suggest that the use of either form of PF technique can lead to improvements in students' writing, with the new form linked to greater improvement. The results of the study also suggest that different effects can occur with learners of different proficiency levels when they are exposed to different forms of PF (see section 3.1.1). Moreover, when each of the writing categories (i.e., mechanics, organisation ...etc.) was tested separately after exposure to each form of PF technique, a statistically significant improvement was found. It is therefore suggested that the use of both forms can have a positive impact on all six categories. The findings also suggest that each form can impact writing positively at both micro and macro levels, with the new form having a greater effect on most of the macro features. Such a finding was to be expected, especially since when using the new form learners were forced to shift their attention to the macro level. This may be because a more concentrated focus on one level can lead to more learning. The finding from the interviews that the effects reported after practice with the new form were related more to macro issues supports this finding.

The interview findings also point to other effects on writing. For instance, two interviewees reported that PF had led them to revise their work before submission. This may indicate that learners were encouraged to assess themselves by revisiting their essays once more before final submission, which suggests that the participants were becoming more self-reliant. In this regard, McMahon (2010) concluded that using self-assessment techniques in conjunction with PF technique can have an even more positive impact on learners.

The questionnaire findings suggest that the students believed that the use of PF technique would affect their writing positively, with a slight positive change in their responses after practice with each form. These minor positive changes may indicate that the students believed in the positive impact of PF, but that this belief was slow to improve after practice with both forms of PF; however, the results of the mid- and post-tests jointly suggest that their writing actually improved significantly at a statistical level. Although the test results suggest a significant positive change in the students' writing, it is possible that the participants needed time in order to be able to sense an actual significant change in their own writing, which explains the significant change in their responses between the pre- and post-questionnaires (i.e., before and after practice with both forms).

Several reasons were reported for the improvement in writing after practice with both forms of PF technique. Additionally, it is possible that the involvement of teacher micro feedback in the new form was a factor in the more significant progress made by the learners in the post-test, which suggests that the students may feel more comfortable when the teacher is involved.

In brief, an examination of the effect PF had on the students' writing quality revealed that both forms had a significant positive effect, taking into account the fact that the writing quality of most of the students was relatively low before they started practising using PF. This finding that PF can lead to an improvement in writing quality is supported by findings from some other studies (Hu, 2005; Lockhart & Ng, 1995; Lundstrom & Baker, 2009; Paulus, 1999; Yang et al., 2006), although it conflicts with the results obtained by Al-Hazemi and Schofield (2007), who reported only a slight improvement in writing. It was also found in the current study that each form of PF led to a significant improvement in all six categories of writing features, with the new form leading to a greater improvement in these categories. As a result, it is suggested that PF led to a deeper improvement in students' writing skills.

## **5. Conclusion**

Ruys and colleagues (2010) and Tolmie and others (2010) argue that collaborative learning can improve the conceptual grasp and application of skills. Tolmie and colleagues (ibid) suggest that this is particularly important because the context in which collaborative learning is utilised can become more positive, and because learning tension is reduced as a consequence of the increase in mutual understanding between learning parties. The findings of our study showed that peer learning can be highly effective, and that it can lead to improvements being made. Additionally, it was found that the new form of PF led to a more positive effect on macro features than on micro features, which may indicate that macro features were not prioritised or even nurtured at all in their past learning.

Macro level features of writing can be neglected by both teachers and students, and focusing the students on these features alone in PF sessions can lead to significant progress in writing. Learners in an L2 context are likely to have an interest in mastering micro writing features, which is likely to lead them to neglect the macro features (Chenoweth & Hayes, 2001; Van Steendam et al., 2010; Van Steendam et al., 2014) that are considered important in the advanced-level English academic programmes in higher education.

The results of the study suggest that the new form is more suitable for high and mid-level learners, but that its suitability can vary in terms of low level students. The evidence did not, however, point to any particular reasons for the contrast in the achievement of low level students, and this matter was in fact beyond the scope of the study and is an area that can be looked at in the future.

Finally, an important point to be mentioned here is that the positive change in learners' writing may have resulted from the different variables involved. Some of these variables are: the fact that the novelty of introducing new approaches to teaching writing might have increased the participants' motivation to learn; the fact that the participants were offered an

increased amount of feedback, with teacher micro-feedback being offered during the second phase; and/or the fact that they practised writing extensively throughout the data collection period. However, it was not possible to determine which factors impacted the participants' writing scores, or the nature of this impact, owing to the design of the study that did not allow for examining the impact of each treatment in isolation.

### Acknowledgement

We would like to express our thanks the Faculty of Arts Research Centre at King Saud University for its invaluable support and encouragement to conduct this research study.

### References

- Al-Hazmi, S., & Scholfield, P. (2007). Enforced revision with checklist and peer feedback in EFL writing: The example of Saudi university students. *Scientific Journal of King Faisal University (Humanities and Management Sciences)*, 8(2), 237-267.
- Alyousef, H. S., & Picard, M. Y. (2011). Cooperative or collaborative literacy practices: Mapping metadiscourse in a business students' wiki group project. *Australasian Journal of Educational Technology*, 27(3), 463-480.
- Broekkamp, H., & Van den Bergh, H. (1996). Attention strategies in revising foreign language text. In G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Eds.), *Theories, models and methodology in writing research* (pp. 170-181). Amsterdam: Amsterdam University Press.
- Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing generating text in L1 and L2. *Written Communication*, 18(1), 80-98. doi: <http://dx.doi.org/10.1177/0741088301018001004>
- Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, 98(4), 891-901. doi: <http://dx.doi.org/10.1037/0022-0663.98.4.891>
- Ferris, D. R. (2004). The “grammar correction” debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime...?). *Journal of Second Language Writing*, 13(1), 49-62. doi: <https://dx.doi.org/10.1016/j.jslw.2004.04.005>
- Gielen, S., Tops, L., Dochy, F., Onghena, P., & Smeets, S. (2010). A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum. *British Educational Research Journal*, 36(1), 143-162. doi: <https://dx.doi.org/10.1080/01411920902894070>
- Gray, C. D., & Kinnear, P. R. (2012). *IBM SPSS statistics 19 made simple*. New York: Psychology Press.
- Hu, G. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research*, 9(3), 321-342.
- 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. doi: <https://dx.doi.org/10.1007/s11251-008-9086-1>
- Hyland, K., & Hyland, F. (2006). Contexts and issues in feedback on L2 writing: An introduction. *Feedback in second language writing: Contexts and issues*, 1-19.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*: Routledge.
- Leki, I. (1991). The preferences of ESL students for error correction in college level writing classes. *Foreign Language Annals*, 24(3), 203-218. doi: <https://dx.doi.org/10.1111/j.1944-9720.1991.tb00464.x>
- Liu, N.-F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279-290. doi: <https://dx.doi.org/10.1080/13562510600680582>
- Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language learning*, 45(4), 605-651. doi: <https://dx.doi.org/10.1111/j.1467-1770.1995.tb00456.x>
- Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18(1), 30-43. doi: <https://dx.doi.org/10.1016/j.jslw.2008.06.002>
- McMahon, T. (2010). Peer feedback in an undergraduate programme: Using action research to overcome students' reluctance to criticise. *Educational Action Research*, 18(2), 273-287. doi: <https://dx.doi.org/10.1080/09650791003741814>
- Min, H.-T. (2005). Training students to become successful peer reviewers. *System*, 33(2), 293-308. doi: <https://dx.doi.org/10.1016/j.system.2004.11.003>
- Min, H.-T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15(2), 118-141. doi: <https://dx.doi.org/10.1016/j.jslw.2006.01.003>
- Min, H.-T. (2008). Reviewer stances and writer perceptions in EFL peer review training. *English for Specific Purposes*, 27(3), 285-305. doi: <https://dx.doi.org/10.1016/j.esp.2008.02.002>
- Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education*, 27(4), 309-323. doi: <https://dx.doi.org/10.1080/0260293022000001337>
- Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8(3), 265-289. doi: [https://dx.doi.org/10.1016/S1060-3743\(99\)80117-9](https://dx.doi.org/10.1016/S1060-3743(99)80117-9)
- Ruys, I., Van Keer, H., & Aelterman, A. (2010). Collaborative learning in pre - service

teacher education: an exploratory study on related conceptions, self - efficacy and implementation. *Educational Studies*, 36(5), 537-553. doi: <https://dx.doi.org/10.1080/03055691003729021>

Strijbos, J.-W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency? *Learning and Instruction*, 20(4), 291-303. doi: <https://dx.doi.org/10.1016/j.learninstruc.2009.08.008>

Tolmie, A. K., Topping, K. J., Christie, D., Donaldson, C., Howe, C., Jessiman, E., et al. (2010). Social effects of collaborative learning in primary schools. *Learning and Instruction*, 20(3), 177-191. doi: <https://dx.doi.org/10.1016/j.learninstruc.2009.01.005>

Truscott, J. (1999). The case for “The case against grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing*, 8(2), 111-122. doi: [https://dx.doi.org/10.1016/S1060-3743\(99\)80124-6](https://dx.doi.org/10.1016/S1060-3743(99)80124-6)

Tsui, A., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147-170. doi: [https://dx.doi.org/10.1016/S1060-3743\(00\)00022-9](https://dx.doi.org/10.1016/S1060-3743(00)00022-9)

Van der Pol, J., Van den Berg, B. A. M., Admiraal, W. F., & Simons, P. R. J. (2008). The nature, reception, and use of online peer feedback in higher education. *Computers & Education*, 51(4), 1804-1817. doi: <https://dx.doi.org/10.1016/j.compedu.2008.06.001>

Van Steendam, E., Rijlaarsdam, G., Sercu, L., & Van den Bergh, H. (2010). The effect of instruction type and dyadic or individual emulation on the quality of higher-order peer feedback in EFL. *Learning and Instruction*, 20(4), 316-327. doi: <https://dx.doi.org/10.1016/j.learninstruc.2009.08.009>

Van Steendam, E., Rijlaarsdam, G., Van den Bergh, H., & Sercu, L. (2014). The mediating effect of instruction on pair composition in L2 revision and writing. *Instructional Science*, 42(3), 1-23. doi: <https://dx.doi.org/10.1007/s11251-014-9318-5>

Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15(3), 179-200. doi: <https://dx.doi.org/10.1016/j.jslw.2006.09.004>

Zhao, H. (2010). Investigating learners' use and understanding of peer and teacher feedback on writing: A comparative study in a Chinese English writing classroom. *Assessing Writing*, 15(1), 3-17. doi: <https://dx.doi.org/10.1177/1362168811426248>

Zheng, C. (2012). Understanding the learning process of peer feedback activity: An ethnographic study of exploratory practice. *Language Teaching Research*, 16(1), 109-126.



## Appendix A

**Table 1: Paired Samples Test (Linguistic Progress)**

		Paired Differences					t	df	Sig. (2-tailed)
		95% Confidence Interval of the Difference							
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	pretest - midtest	-1.09756	.62581	.09773	-1.29509	-.90003	-11.230	40	.000
Pair 2	midtest - posttest	-1.61992	.85087	.13288	-1.88849	-1.35135	-12.191	40	.000
Pair 3	DIFF1 - DIFF2	.52236	1.20157	.18765	.14310	.90162	2.784	40	.008

**Table 2: Paired Samples Statistics (effects on writing features: macro & micro)**

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	PRE_Organisation	3.5244	41	1.31316	.20508
	MID_Organisation	4.8537	41	1.02007	.15931
Pair 2	MID_Organisation	4.8537	41	1.02007	.15931
	POST_Organisation	6.5122	41	1.27714	.19946
Pair 3	diff1Organisation	-1.3293	41	.85593	.13367
	diff2Organisation	-1.6585	41	.83246	.13001
Pair 4	PRE_Development	3.4024	41	1.17909	.18414
	MID_Development	4.5488	41	.99250	.15500
Pair 5	MID_Development	4.5488	41	.99250	.15500
	POST_Development	6.3415	41	1.26214	.19711
Pair 6	diff1Development	-1.1463	41	.91681	.14318
	diff2Development	-1.7927	41	1.02455	.16001
Pair 7	PRE_Cohesion	3.2805	41	1.28476	.20065
	MID_Cohesion	4.4756	41	1.26467	.19751
Pair 8	MID_Cohesion	4.4756	41	1.26467	.19751
	POST_Cohesion	6.4024	41	1.21047	.18904

Pair 9	diff1Cohesion	-1.1951-	41	.98665	.15409
	diff2Cohesion	-1.9268-	41	1.06396	.16616
Pair 10	Pre_Structure	3.2805	41	1.40111	.21882
	Mid_Structure	4.1341	41	1.11817	.17463
Pair 11	Mid_Structure	4.1341	41	1.11817	.17463
	Post_Structure	5.6707	41	1.48581	.23204
Pair 12	diff1Structure	-.8537-	41	.88207	.13776
	diff2Structure	-1.5366-	41	1.18528	.18511
Pair 13	Pre_Vocabulary	3.2683	41	1.22013	.19055
	Mid_Vocabulary	4.3902	41	1.09836	.17154
Pair 14	Mid_Vocabulary	4.3902	41	1.09836	.17154
	Post_Vocabulary	5.9024	41	1.30489	.20379
Pair 15	diff1Vocabulary	-1.1220-	41	.79672	.12443
	diff2Vocabulary	-1.5122-	41	1.02767	.16049
Pair 16	Pre_Mechanics	3.8293	41	1.23293	.19255
	Mid_Mechanics	4.7683	41	.96888	.15131
Pair 17	Mid_Mechanics	4.7683	41	.96888	.15131
	Post_Mechanics	6.0610	41	1.36563	.21328
Pair 18	diff1Mechanics	-.9390-	41	.80774	.12615
	diff2Mechanics	-1.2927-	41	1.01843	.15905

**Table 3: Paired Samples Test (effects on writing features: macro & micro)**

		Paired Differences							
				95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	PRE_Organisation - MID_Organisation	-1.32927-	.85593	.13367	-1.59943-	-1.05910-	-9.944-	40	.000
Pair 2	MID_Organisation - POST_Organisation	-1.65854-	.83246	.13001	-1.92129-	-1.39578-	-12.757-	40	.000
Pair 3	diff1Organisation - diff2Organisation	.32927	1.27284	.19878	-.07249-	.73103	1.656	40	.105
Pair 4	PRE_Development - MID_Development	-1.14634-	.91681	.14318	-1.43572-	-.85696-	-8.006-	40	.000

Pair 5	MID_Development POST_Development	-1.79268-	1.02455	.16001	-2.11607-	-1.46930-	-11.204-	40	.000
Pair 6	diff1Development diff2Development	.64634	1.55822	.24335	.15451	1.13818	2.656	40	.011
Pair 7	PRE_Cohesion MID_Cohesion	-1.19512-	.98665	.15409	-1.50655-	-.88370-	-7.756-	40	.000
Pair 8	MID_Cohesion POST_Cohesion	-1.92683-	1.06396	.16616	-2.26266-	-1.59100-	-11.596-	40	.000
Pair 9	diff1Cohesion diff2Cohesion	.73171	1.69962	.26544	.19524	1.26817	2.757	40	.009
Pair 10	Pre_Structure Mid_Structure	-.85366-	.88207	.13776	-1.13207-	-.57524-	-6.197-	40	.000
Pair 11	Mid_Structure Post_Structure	-1.53659-	1.18528	.18511	-1.91070-	-1.16247-	-8.301-	40	.000
Pair 12	diff1Structure diff2Structure	.68293	1.69468	.26467	.14802	1.21783	2.580	40	.014
Pair 13	Pre_Vocabulary Mid_Vocabulary	-1.12195-	.79672	.12443	-1.37343-	-.87048-	-9.017-	40	.000
Pair 14	Mid_Vocabulary Post_Vocabulary	-1.51220-	1.02767	.16049	-1.83657-	-1.18782-	-9.422-	40	.000
Pair 15	diff1Vocabulary diff2Vocabulary	.39024	1.55930	.24352	-.10193-	.88242	1.603	40	.117
Pair 16	Pre_Mechanics Mid_Mechanics	-.93902-	.80774	.12615	-1.19398-	-.68407-	-7.444-	40	.000
Pair 17	Mid_Mechanics Post_Mechanics	-1.29268-	1.01843	.15905	-1.61414-	-.97123-	-8.127-	40	.000
Pair 18	diff1Mechanics diff2Mechanics	.35366	1.33811	.20898	-.06870-	.77602	1.692	40	.098

**Table 4: Paired Samples Statistics (comparison of development in macro and micro features)**

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	PreMicrofeatures	3.4593	41	1.22008	.19054
	PreMacrofeatures	3.4024	41	1.16068	.18127
Pair 2	MidMicrofeatures	4.4309	41	.98811	.15432
	MidMacrofeatures	4.6260	41	1.03265	.16127
Pair 3	PostMicrofeatures	5.8780	41	1.30119	.20321

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	PreMicrofeatures	3.4593	41	1.22008	.19054
	PreMacrofeatures	3.4024	41	1.16068	.18127
Pair 2	MidMicrofeatures	4.4309	41	.98811	.15432
	MidMacrofeatures	4.6260	41	1.03265	.16127
Pair 3	PostMicrofeatures	5.8780	41	1.30119	.20321
	PostMacrofeatures	6.4187	41	1.18035	.18434

**Table 5: Paired Samples Test (comparison of development in macro and micro features)**

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	PreMicrofeatures - PreMacrofeatures	-.05691	.58997	.09214	-.12931-	.24313	.618	40	.540
	MidMicrofeatures - MidMacrofeatures	-.19512	.43764	.06835	-.33326-	-.05698-	-2.855-	40	.007
Pair 3	PostMicrofeatures - PostMacrofeatures	-.54065	.49129	.07673	-.69572-	-.38558-	-7.047-	40	.000

**Table 6: Test Statistics<sup>i</sup> (effects on writing)**

	MidPFeffectonwriting - PrePFeffectonwriting	PostPFeffectonwriting - MidPFeffectonwriting
Z	-1.687 <sup>ii</sup>	-.769 <sup>a</sup>
Asymp. Sig. (2-tailed)	.092	.442

**Table 7: Test Statistics<sup>iii</sup>(effects on writing)**

	PostPFeffectonwriting - PrePFeffectonwriting
Z	-2.444 <sup>iv</sup>
Asymp. Sig. (2-tailed)	.015

**Copyright Disclaimer**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).

---

<sup>i</sup>Wilcoxon Signed Ranks Test.

<sup>ii</sup>Based on negative ranks.

<sup>iii</sup>Wilcoxon Signed Ranks Test.

<sup>iv</sup>Based on negative ranks.