

Effectiveness of Peer Assessment and Peer Feedback in Pakistani Context: A Case of University of the Punjab

Nadia Majid* and Muhammad Islam**

Abstract

Writing an effective essay in English is a serious challenge for most of the students, despite its academic importance, in Pakistan. The methodology used to teach writing skills is ineffective and the assessment criteria is outdated. In this regard, peer assessment and peer feedback practices have proved to be fruitful in enhancing the writing skills of second language learners. Therefore, this study aimed to examine whether there is any improvement in the overall writing skills as well as the specific subskills of writing after implementing peer assessment and peer feedback practices. A quasi-experimental design, called Multiple Baseline Time Series design, was used to collect data from the students during five interventions of essay writing, peer assessment and peer feedback process. Two intact groups of undergraduate students were chosen, through cluster random sampling, as experimental and control groups, from Center for Undergraduate Studies at University of the Punjab. The data collected, was in the form of essays written by the students and the scores given by the peers during peer assessment. Those scores as well as content analysis of the essays helped the researcher to analyze the results of the study. The findings reveal that although there was a slight improvement in a few subskills of writing, yet in the Pakistani context, these assessment practices failed to present any substantial learning achievement overall, as compared to traditional teacher-based assessment practices. Therefore, the study concluded that in our context, these assessment practices may be productive, if they are implemented with teacher's assessment, after thorough training of the teachers as well as students.

Keywords: Peer assessment, peer feedback, essay writing, second language writing skills, quasi-experimental study, content analysis.

*College of Art and Design, University of the Punjab, Lahore Email: nadia.majid29@hotmail.com

**Assistant Professor, Institute of Education and Research, University of the Punjab, Lahore

Introduction

Peer assessment and peer feedback practices have been a crucial part of educational settings for more than three decades. However, the recent trends in assessment have attached a significance to these assessment practices, in promoting learning. For instance, there is a great emphasis on goal-oriented learning to encourage lifelong learning in the students. The movement 'Assessment for Learning' has become widespread which emphasizes the importance of engaging students in the assessment practices and involving them in feedback processes. Zhang (2018) explains that there is an urgent need for a more effective approach, through which students receive richer feedback, engage in bidirectional communications and advance their writing as well as learning capabilities. In this regard, peer assessment and peer feedback have gained relative popularity, in enhancing the writing skills. Literature provides substantial evidence of peer assessment and peer feedback proving to be beneficial in improving the writing skills of students in various contexts in studies conducted across the globe (Gencha, 2018; Kuyyogsuy, 2019; Meletiadiou, 2021a; Uymaz, 2019; Wu & Schunn, 2021; Zhang, 2018).

Topping (2017) defines peer assessment as an arrangement for learners to consider and specify the quality of a product, of other equal-status learners, which leads to learning further, by giving elaborated feedback to achieve a negotiated agreed outcome. In other words, peer assessment is a valuable pedagogical practice as it enables the learners to take part in assessment by evaluating their peers' learning process and products (Bryan & Clegg, 2019). In addition, analysing other students' work allows students to interpret the assessment criteria, which leads to an understanding of good performance and helping students to adjust their actions to meet the expected results (Ion, Marti & Morell, 2019). Peer assessment incorporating peer feedback leads to more beneficial outcomes as peer feedback provides the strengths and weaknesses along with recommendations for improvement (Meletiadiou, 2021a). Furthermore, peer feedback means to have a dialogue whereby students share knowledge and understanding with the intention of informing as ongoing learning (Zhu & Carless, 2018).

Peer assessment has gained relative popularity in higher education owing to its favorable outcome in enhancing the writing skills, developing the critical ability, promoting collaboration, improving higher order skills and taking responsibility of one's own learning. Detailed studies of peer assessment have shown that students benefit from it (Double et. al. 2020; Meletiadiou, 2021b). The learners' perspectives about the benefits of peer assessment have also been explored and have turned out to be favourable in most of the studies (Gencha, 2018). Students believe that peer assessment and peer feedback facilitate learning and motivation (Gencha, 2018; Quynh, 2021; Wu & Schunn, 2021). Huisman et. al. (2019) explain that the benefit of peer feedback is that it is available in

greater volume and is given with greater immediacy, as compared to teacher feedback. Since it is quite time consuming for teachers to provide detailed feedback, peer feedback plays a more eminent role in students' learning.

However, peer assessment and feedback have not been able to escape critique. Various limitations have also been reported regarding these assessment practices in literature. Researchers have found that although students express positive attitudes toward the usage of peer feedback, they tend to significantly favor the feedback given by the teachers (Quynh, 2021). There is also a possibility that students' negative attitudes towards peer review might lead them to be unfair with their peers. In this way, they may provide over-critical comments about their peers' writings (Rouhi & Azizian, 2013). Moreover, at the initial stage, students tend to be more doubtful of their peer's ability to assess their work (Ashenafi, 2017).

Consequently, literature provides recommendations to make the most of peer assessment and feedback to promote learning. It is agreed that students need support in order to provide quality feedback (Walker, 2015). This support may be provided by the teacher in the form of training of the students for peer assessment, preferably by using a criterion (Topping, 2009). Wu & Schunn (2021) suggest that by implementing anonymous peer feedback to decrease bias, use of well-structured rubrics for peer assessment and sufficient training in giving feedback, these assessment practices become more reliable and profitable. Hence, despite all the critique, peer assessment and peer feedback have been able to show valuable results.

An important aspect of peer assessment and feedback practices is the link of these activities to Vygotsky's (1978) social development theory, which emphasizes the vital role of social interaction in learning (Lundstorm & Baker, 2009; Topping, 2021). This theory emphasizes the importance of learning through interaction with the society and more specifically in an educational context, the collaboration of students with their teachers and peers. Further, it is claimed that the peer assessment process naturally constructs a favorable teaching environment for peers to work within the zone of proximal development (ZPD) (De Guerrero & Villamil, 2000). The learner's ZPD refers to the place between where learners are able to perform a task on their own versus with the help of a teacher or parent (Lundstorm & Baker, 2009). The theoretical framework of Vygotsky's social development theory refers to two important aspects of peer assessment and peer feedback; *cognitive development* and learning through social interaction which can be implied as *collaborative learning*.

According to one of the first advocates of peer assessment, cognitive and meta-cognitive benefits can accrue before, during, or after the peer assessment (Topping, 2009). The learners critically evaluate the written text during peer assessment and suggest various improvements for the sake of the revision of the text during peer feedback. Hill (2016) claims that feedback triggers cognitive processes which can encourage the verification or adjustment of understanding, point to filling in gaps and indicate alternative strategies. This cognitive development occurs in a social environment which create opportunities of interactive activities promoting learning. Peer assessment and peer feedback provides the learners with exactly those interactive activities which propagates learning from one another instead of learning in a teacher centered classroom.

Topping (1998) asserts that the need to communicate the assessment to another should create purpose and accountability, and the language used for this purpose should lead to the assessor's internal thought processes. Zhang (2018) suggested that in opposition to the traditional class, peer collaboration encouraged students to discover their intended meanings and to express themselves more confidently through their writing. Furthermore, the process of discussion and elaboration during feedback allows the students to work collaboratively to assess each other's work (Gencha, 2018). It also lowers the inhibitions of the students, and they feel less uncomfortable when their errors are being pointed out and they focus more on learning and negotiating meaning through the collaboration.

In Pakistan, writing skills have an essential role at the academic level, both in learning and assessment process, as our primary mode of assessment is the written examination. The student's performance is evaluated through grades achieved on paper pencil tests. Hence, it is crucial to have command on writing skills to succeed. Ironically, the students gain exceptional grades despite having poor writing skills. Their key to success is rote learning and cramming (Akram, 2017; Siddiqui, 2007). Due to this factor, our students struggle to excel in writing skills and even after reaching the graduate level, they are incapable of expressing their thoughts in a creative and coherent way.

Usually in our classrooms, during essay writing sessions, the teacher introduces the topic to the students and then provides them with a sample essay to learn and replicate (Rahman, 2010). Most of the times, the students are discouraged to write creatively on their own because it increases the burden of checking and making corrections for the teacher. Nevertheless, when the students are presented with an opportunity to write on their own, they find it difficult to come up with ideas and correct sentence structure as well as appropriate vocabulary to write effectively. When it comes to assessment of these essays, the teachers mostly point out the spelling and grammar mistakes (Haider, 2012).

In large classroom settings, feedback provided by the teacher is mostly insufficient or delayed which does not contribute much to the improvement of English language writing skills. To grade an essay written by the students, instead of using some specific criteria of assessment or a rubric, the teacher gives a general score, along with a few red underlined marks and some spelling or grammar errors, encircled in the essay. These circumstances describe the need of more practical and productive methods of teaching and assessment, to ensure learning achievement in the students. This leads to the statement of the problem of the present study.

Statement of the Problem

Previous studies in Pakistani context as well as common observation and experience of professionals in education, indicates that a large number of students fails in English due to poor writing skills (Siddique & Singh, 2017) yet no serious efforts are usually made to address the issue (Sultana & Zaki, 2015). The education system in Pakistan, still follows the traditional method of teaching and assessment of English (Akram, 2017) due to which the writing skills of most of the learners do not seem to substantially improve, even when they reach the tertiary level of education (Fareed, Jamal & Zai, 2021). They study English as a compulsory subject for at least twelve years before they reach undergraduate level, yet they are incapable of communicating in both written and spoken English (Sarfraz, Mansoor & Tariq, 2015).

Moreover, the curriculum developers know that English language testing does not measure students' creativity and critical thinking, still there is no significant efforts being made to improvise the teaching and testing methodology (Khan, 2011). Reliance on outdated methods of teaching and assessment, may be one of the reasons behind the weak English writing skills. Under these circumstances, it is evident that alternative feedback practices that are effective and practically efficient, are required (Huisman et. al., 2019). Akhtar et. al., (2019) emphasize that it is essential to introduce such approaches which may focus on collaborative learning among students. Therefore, there is a dire need to introduce effective teaching methodology and investigate alternative methods of assessment, which encourages the development of writing skills in Pakistan.

Various studies on peer assessment and peer feedback concerning the development of English language writing skills have been conducted all over the world, yet the benefits of peer assessment and peer feedback in improving the English language writing skills of undergraduate students in Pakistan had not been thoroughly investigated. Hence, this study aimed to fill this gap and offered an insight into the effectiveness of peer assessment and peer feedback in a Pakistani context.

Significance of the Study

This study was significant in its own way, as such an experimental study, had not been reported so far, which concerned the effect of peer assessment and peer feedback practices on the undergraduate students' English language writing skills in Pakistan, specifically in the University of the Punjab. Furthermore, in Pakistan, there is a dearth of studies which explore the interaction of students among themselves to collaboratively contribute to the development of their writing skills in English. This is important because in the higher-level English classrooms, the large number of students make it impossible for the teacher to give extensive and individual feedback to students, on their essays.

Moreover, the results of the present study are significant in highlighting the importance of peer assessment and peer feedback in developing the critical skills of the learners. This was due to the responsibility of assessment bestowed upon the students as well as the dialogue sessions during peer feedback. The students were compelled to use their analytical ability to develop a critical outlook during peer assessment and feedback practices. Consequently, it may reduce the students' teacher dependence and encourage them to share the teacher's workload of assessment.

In addition, this study may prove to be a groundwork for successful implementation of peer assessment and feedback practices in future in Pakistan, to upgrade the writing skills of the students. These assessment practices have proved to be effective in most of the second language learning settings and with proper implementation in our context, it may prove to be fruitful. Further, this study proposed measures to be taken, for successful implementation of peer assessment and peer feedback practices. It may prove to be a way forward to bring a change in the teaching, learning and assessment of writing skills of the students.

Objectives of the Study

The objectives of the study investigate whether peer assessment and peer feedback can improve the English language writing skills of the students. Furthermore, the study also examines that if the writing skills of the students are enhanced, then which subskills of writing, may improve more than others.

Research Questions

1. Are the students able to learn and improve their writing skills through peer assessment and peer feedback?
2. Which sub-skills of English language writing skills are improved more, as compared to others, owing to peer assessment and peer feedback?

Methodology

To determine the effect of peer assessment and peer feedback on the writing skills of students, it was essential to design the study in a way where students were practically made to write the essays, assess them and give feedback to their peers. Based on the theoretical underpinnings of the study, a conceptual framework was developed to evaluate the extent of improvement in the writing skills of the students owing to peer assessment and feedback practices. The present study used Multiple Baseline Time Series design as it allows a repetition of interventions, where it is impossible to reverse to the baseline condition (Fraenkel & Wallen, 2009). Writing skills is such a variable that a change in it through peer assessment and feedback could not be reversed and every intervention of essay writing task only added to the behavior of the student. Hence, this longitudinal design helped in examining the trends in the data, at multiple time points, before, during, and after an intervention (Marczyk, DeMatteo & Festinger, 2005).

In experimental studies, threats to external and internal validity are unavoidable. To ensure external validity, several pre-requisites were defined about the sample of students used in the study. For instance, the students were enrolled in a Bachelor Honours degree programme at the University of the Punjab. They had a background of previously studying in public or private institutes due to which they had a diverse social and linguistic background and if any other group matches these pre-requisites, the study may be generalized.

Furthermore, to explain the internal validity of Multiple Baseline Time Series Design, Gay, et al. (2012) have given a profound explanation of how the use of this design eliminates the two major internal threats of 'history' and 'instrumentation'. History may pose as a problem in time series design because some event or activity may occur between the last pretest and the first posttest. Therefore, in this experimental study the presence of a control group made it possible to observe any changes in the two groups due to any unavoidable circumstances as the difference would be evident in the results. The second threat of instrumentation may appear to be a threat if only the researcher changed the measuring instruments during the study (Gay, et. al. 2012). This did not pose as a problem in this study as the instruments were kept the same during the entire experimental study.

In addition, Multiple time series design essentially involves the addition of a control group to the basic design and this eliminates instrumentation as validity threats (Gay, et al. 2012). Hence, a control group resolves the issues of internal validity in a quasi-experiment because the researcher has little or no control over the quasi-independent variable and other related variables (Bordens & Abbott, 2018). Furthermore, the presence of a control group allows to assess whether any post treatment changes in performance would have happened anyway if the treatment had not been introduced at that time (Bordens & Abbott, 2018).

In this study, the target population was the undergraduate students enrolled in the Bachelor programme at the Center for Undergraduate Studies, University of the Punjab. Two intact groups were selected through cluster random sampling, out of the five available intact Bachelor classes. It was crucial to gain access to the students in a scheduled class, as they were required to write essays and then peer assess them. Therefore, the intact class of students selected as the experimental group had 27 students while the control group had 22 students. The data was collected from the two groups simultaneously over a period of one semester (almost 14 weeks) following the procedure shown in the figure below.

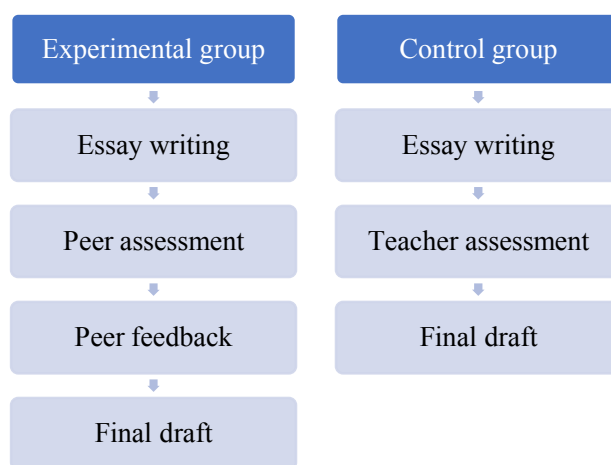


Figure 1. *Data collection procedure*

As illustrated in the figure above, the data collection procedure for the students of experimental group, required them to write an essay in the class. Those essays were photocopied and allotted a fictitious number for anonymity. Each student was given an essay to peer assess based on a rubric. The students were asked to give written feedback along with oral feedback during the feedback sessions. After the feedback, the students were asked to write a final draft of the essay. This intervention was repeated five times on alternate weeks.

On the other hand, the students of the control group were asked to write an essay, in class, on the same topics as the experimental group. However, their essays were assessed by the teacher against the same rubric that was used by the students for assessment. These students also wrote a final draft after the teacher's assessment and feedback. To ensure that the students write a creative essay without any role of rote learning, they were given the topic right before the activity in class, and it was different from the common topics attempted in the board examinations.

The instrument used to collect the quantitative data from the students, was through the rubric (see Appendix A) which was used to assign scores by the students, during peer assessment. The rationale of using a rubric in the present study was to ensure a transparent process of assessment among the students. It provided the students with a clear guidance about the subskills to be assessed in the essays. In addition, it laid down specific guidelines and descriptions to rank and score the essays which streamlined the process of assessment. It served as a scaffold for the students, especially for the low ability students who could have found peer assessment a daunting task.

The rubric was adapted by the researchers, from a study by Lundstorm and Baker (2009), who used it as a rubric to grade essays during peer review activity by the students. The reason of using this particular rubric is that it covers the global aspects of writing skills, for instance organization, cohesion, grammar, vocabulary, spelling and punctuation etc. To determine the inter-rater reliability, two expert teachers of English language with more than ten years' experience of teaching English essay writing were asked to rate three different essays using the rubric. The scores on each essay by the two raters were averaged. Since the two scores by each expert varied by one point, the average score was accepted as the final score. Similarly, to ensure the validity of the rubric, it was shared with three English language experts who reviewed the rubric and agreed it to be valid, after the suggested changes by them.

Effect on Writing Skills through Peer Assessment and Peer Feedback

The data collected during the interventions was analyzed through different analysis to evaluate the effect on the writing skills of the students. First, the scores of the essays, attained by the students were analyzed through One- way repeated measures ANOVA to determine the overall improvement in the writing skills of both the experimental and control group. Then to evaluate which subskills of English language improved more according to the students' evaluation during peer assessment, paired samples t-test was administered. It is noteworthy to mention here that the scores attained by the students were only used to determine whether they were capable of objectively scoring essays and whether it helped in their cognitive development. Hence, the students were clarified that the scores would not be used for any summative or formative assessment that could influence their grading of the subject English that they were studying in their undergraduate programme.

Furthermore, since the improvement in the writing skills of the students could not only be measured through the peer assessment scores of the students, hence the researcher also analyzed the students' essays through content analysis to evaluate whether there was any substantial effect on the students' writing. Table 1 reports the performance of the experimental and control group students in all the five interventions of essay writing activity. One-way repeated measures ANOVA was used to analyze the scores attained by the students against the rubric provided to them for assessment.

Table 1

Overall mean scores of all the five essays of experimental group and control group

	N	Essay 1 Mean	Essay 2 Mean	Essay 3 Mean	Essay 4 Mean	Essay 5 Mean	Sig. $p < 0.05$	Wilks' Lamda	Eta
Experimental Group	27	18.37	19.07	18	19.96	20.56	0.003	0.510	0.490
Control Group	22	13.59	15.73	17.55	18.55	17.59	0.000	0.332	0.668

Table 1 shows the mean values of the five interventions of essay writing for both the experimental and control group. It can be noticed that if we compare the mean of essay 1 with essay 5 for both the groups of students, there is an increase in the mean values. It is evident that both the experimental and control group display statistically significant results as the significance value for both the groups is $p < 0.05$ implying that there was an overall improvement in the writing skills of both the groups. In this study, the control group not only resolved the issues of internal validity of the experiment but also displayed the natural progression in the writing skills of the students in a traditional classroom.

It is noteworthy that the students of the experimental group attained scores on their essays through peer assessment while the students of the control group were scored by the teacher, against the same rubric used for peer assessment. Yet, it would be unfair to compare the findings of two groups as they were assessed by different assessors (peers and teacher) of which, peer assessors were not even equally qualified for assessment. Hence, it was difficult to evaluate whether the students' assessment was reliable to be accepted as true evaluation of the essays and reflective of actual improvement.

Therefore, to determine whether the students of the experimental group displayed any improvement in their writing skills, the researcher conducted two different analyses. The first one was based on the scores given by the students to their peers against the rubric. While the second analysis was conducted by the same teacher who assessed the students of the control group in the form of content analysis of the essays written by the students of the experimental group. The results of the both the analyses are presented in the section below.

Effect of Peer Assessment and Feedback on the subskills of English writing skills

To evaluate which subskills of English language writing skills improved more, according to the students' evaluation during peer assessment, paired samples t-test was administered and the mean scores of the first essay was compared to the mean scores of the fifth essay.

Table 2

Comparison of the first and last essay scores of peer assessment of experimental group

	Essay 1		Essay 5		Mean Difference	t	Sig.(2-tailed)	Eta
	Mean	SD	Mean	SD				
Content	3.07	1.07	3.52	0.89	-0.444	-2.000	0.056	0.143
Organization &unity	3.00	0.88	3.52	0.89	-0.519	-2.331	0.028	0.173
Coherence & Cohesion	3.22	0.85	3.26	0.86	-0.037	-0.161	0.873	0.001
Vocabulary	2.81	1.00	3.41	0.89	-0.593	-3.049	0.005	0.263
Grammar	2.89	0.89	3.37	0.97	-0.481	-2.229	0.035	0.160
Spelling& Punctuation	3.37	1.21	3.48	0.85	-0.111	-0.431	0.670	0.007

N=27

This table gives a detailed description of the extent of improvement in each category. The results clearly show that there is a statistically significant increase in the mean scores of three subskills: 'Organization and unity', 'Vocabulary' and 'Grammar'. Hence, it suggests that the students showed significant improvement in these three categories. Apparently, the maximum improvement is witnessed in the sub skill 'vocabulary' with a mean difference of -0.593 in the essay scores. The second significant improvement is noticeable in the sub skill 'organization and unity' with a mean difference of -0.519 in the scores of the essays. The third evident improvement was observed in the sub skill 'grammar' with a statistically significant mean difference of -0.481 in the essay scores. It is interesting to notice that the mean difference in all the categories is a negative value reflecting that there is an overall improvement in all the categories, but only three showed a more statistically significant improvement. It is noteworthy that these findings are only based on the scores provided by the students to their peers hence they may not reflect the true findings. Therefore, the content analysis by the teacher was also considered necessary to determine the improvement in the writing skills of the students.

Content Analysis of the Essays

This study used the framework of deductive content analysis proposed by Mayring (2000) for content analysis depicted in figure 2.

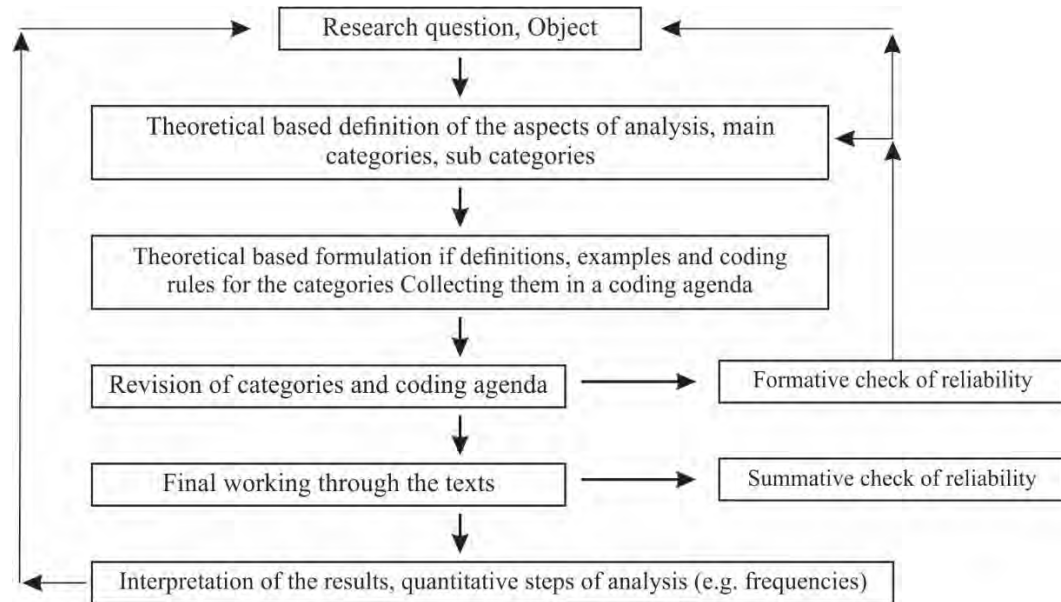


Figure 2. *Model of deductive content analysis (Mayring, 2000)*

Graneheim, Lindgren and Lundman (2017) explain that qualitative content analysis is a method to analyze qualitative data, which concentrates on finding the meaning of the subject, context and emphasizes upon the variations occurring in the text. This is a research method in which data is interpreted through the systematic classification process of coding and identifying themes or patterns. The purpose of conducting a content analysis was particularly to study the improvement in the subskills of writing, hence deductive content analysis was deemed appropriate. Mayring's (2000) content analysis framework illustrates that the research questions lead to the development of the main categories and their descriptions. These descriptions in turn give rise to the coding rules for the categories. The categories are revised, and a quantitative analysis is conducted according to the relevant method. Therefore, the essays were thoroughly analyzed for the process of developing subcategories from the main categories.

There were 27 students in the experimental group and each student wrote 5 essays. Every essay was written in two drafts; the first one, which was also used for peer assessment and peer feedback and then a final draft of the same essay, written after receiving feedback. To determine which aspects of writing skills improved more, the researcher observed the changes in the fifth essay as compared to the first essay.

Therefore, for every student, the first and the fifth essays' first drafts were chosen for content analysis. The reason to choose the first draft of both the first and fifth essay was that those essays were written in the class in the presence of the teacher. On the other hand, the final drafts were always written at home after the students had received peer feedback. Apparently, the students might have sought help from various sources to improve the final draft of the essay and thus they did not reflect the true improvement in their final drafts of their ability to write well. Consequently, 54 essays were selected and used as sample for content analysis.

During the content analysis, the number of errors in the first essay, were compared to the fifth essay. The number of students vary in the first essay according to the students who made errors in the relevant subskill. Those students were omitted in the counting who did not make errors in the relevant subskill in the first essay. Thus, it was evaluated whether the students were able to exhibit any improvement by the fifth essay or not for those who made errors in the first essay. The results are presented in table 3.

Table 3

Comparison of the first and last essay of experimental group through content analysis

Category	Description	Sub category/ Examples of errors	No. of sts. with errors in the 1st essay	No. of sts. who improved in the 5th essay	Percentage of sts. who improved
Content	The content of the essay should be completely relevant to the topic. There should be a thorough development of ideas.	Some of the content was irrelevant. Too general content. Essay was concise and lacked details.	19	7	37%
Organization	The content should be organized into proper paragraphs as introduction, main body and conclusion accordingly.	Essay written in one paragraph only. No proper division of paragraphs. Unorganized ideas in paragraphs.	14	3	21%
Coherence & cohesion	There should be a good progression of ideas with proper transition of paragraphs using discourse markers or topic sentences.	Poor connection between paragraphs. No flow of ideas. No topic sentences. Lack of discourse markers.	16	4	25%
Vocabulary	There should be use of appropriate vocabulary, a variety of words and suitable idiomatic phrases.	Repeated use of the same words. No use of synonyms or idioms. Use of common vocabulary.	24	8	33%
Grammar	There should be correct use of forms of verbs, tense and structure of the sentences.	Incorrect sentence structures. Incorrect use of tense, verb form or parts of speech.	25	6	24%
Spelling	The spelling of the words in the essay should be correct.	Spelling errors.	16	6	38%
Punctuation	There should be correct use of punctuation marks and capitalization.	Lack of punctuation. Incorrect punctuation.	14	3	22%

The table of content analysis reveals the percentage of students who showed improvement in the fifth essay. The results in table 3, paint a different picture because through the Paired samples t-test, a significant improvement was noticed in 'organization', 'vocabulary' and 'grammar'. While, through content analysis the maximum improvement was noticed in 'content', 'vocabulary' and 'spelling'. However, the improvement in each subskill is so minute that we cannot consider peer assessment and feedback effective assessment practices in our context. Therefore, we can conclude that although there was a subtle improvement in each category, but it was evidently not significant.

Discussion

Several studies have been conducted worldwide to ascertain the effectiveness of peer assessment and feedback on second language learning (Gencha, 2018; Quynh, 2021; Wu & Schunn, 2021; Yüce & Aksu Ataç, 2019). Most of the studies showed favorable results despite the limitations involved in the process. In this study, in the Pakistani context, the findings reveal that the students displayed a minor progression in their English language writing skills. Overall, some students showed improvement while for some others, although the scores had increased but the quality of writing did not.

The evidence lies in the findings of the content analysis which does not exhibit much improvement in the writing skills of the students. This reflects that peer assessment and peer feedback did not bring about a substantial improvement in the writing skills of the students. This is also confirmed through the comparison of the results of the experimental group with the control group students. We learn that both the groups performed equally as they both demonstrated a slight improvement in the writing skills. Hence, it may be assumed that peer assessment and feedback practices may have not been able to outperform the traditional method of teacher assessment in Pakistani context.

To gain an insight into the intriguing results of this study, we take into account the achievement of both the groups involved in the experiment. It cannot be denied that the students of the experimental group felt that they had gained more knowledge about the techniques of writing an essay. This was evident from the comparison of the students' performance in their last essay compared to their first essay in the various subskills of writing skills. They believed that they had learnt about the essential aspects involved in writing a good essay.

Interestingly on the other hand, the students in the control group also showed a similar improvement in their essays when their performance was evaluated by comparing the scores of their last essay with their first essay. The students of the control group were not introduced to any interventions, rather their essays were evaluated by the teacher against the same criteria that was used by the students during peer assessment.

These findings are very important as we may imply that in the Pakistani context, peer assessment and peer feedback may not contribute exceptionally to enhance the English language writing skills of the students. Apparently, the reason may be the teacher dependent culture in our education system where the students entirely rely on the teacher in the classroom (Yasmin, et al., 2019). The students of the experimental group did not show a substantial change because they were more dependent on the teacher's assessment rather than peer assessment scores. Most of the students were also unsatisfied with peer feedback and complained that it was insufficient. While the students of the control group also displayed no substantial improvement in their writing skills. The reason may be the teacher's inability to provide extensive feedback to the students.

In Pakistan, mostly there is a large number of students in an undergraduate class of compulsory English subject. This is emphasized by Jokhio, Raza, Younus and Soomro (2020), who maintains that in Pakistani public sector universities' context, English language teaching and learning takes place in large-sized classes (Akram, 2017; Khurram, 2018). It becomes almost impossible for the teachers to invest time in providing the students with extensive feedback on a creative writing assignment. They mostly assign grades to the students and generally discuss the errors made by most of the students in the class. Consequently, there is no improvement in the writing skills of the students.

However, we cannot overlook a slight progression in experimental group students' scores of the subskills of writing that emerged in the findings of this study. The overall findings of quantitative analysis provide evidence that the students improved in 'organization and unity', 'vocabulary' and 'grammar'. While the results of the content analysis by the teacher, show a subtle improvement in 'content', 'vocabulary' and 'spelling'. Overall, both analyses display a development in vocabulary after the interventions. The reason may be that reading others' essays may have introduced them to new words and this exposure enhanced their vocabulary. Most probably, they might have incorporated the newly learnt vocabulary in their own essays. Similarly, the students learnt from reading others' essays that how their peers generated and organized their ideas into different paragraphs.

Furthermore, in Pakistan, English language is largely taught through Grammar Translation Method (Akram, 2017; Rahman, 2010), hence, grammar is the most important aspect of writing skills during teaching and assessment. Therefore, the students also paid particular attention to identifying grammatical errors in their peer's essays during peer assessment and feedback. They assumed that they had improved their grammar after the interventions because of peer assessment. Yet, we learn through the content analysis that the students reduced the errors, but they could not learn the use of correct form of verb and tense structure. The reason might be the absence of a teacher's assistance and feedback as students rely heavily on their teachers' guidance.

Furthermore, the slight improvement in the writing skills of the students may also be because the students had discussions with their peers about the correction of their errors. They cross checked it with the teacher or the internet to clarify their confusions. This dialogue between the student and peer assessor not only proved beneficial but also developed a sense of critical thinking in the students. Zheng, Cui, Li and Huang (2018) explain that when assessors and assessee are exposed to conflicting ideas, synchronous group discussion can fill the gaps in their understanding.

In addition, the slight improvement in the subskills of writing may also be attributed to the use of rubric against which the essays were scored in this study. It is in line with the findings of Wang (2014), where it was suggested that rubrics may be more widely used for facilitating students' peer feedback. The rubric consisted of various subskills and their descriptors which provided the students with a better understanding of what they should focus on while writing an essay. Usually, the English teachers in Pakistan do not use any particular criteria or rubric to assess the essays because of which the students have never been truly able to understand the scoring process. Hence, the use of rubric to assess the essays in this study kept them focused on the subskills that were being particularly focused on during the marking process.

To sum up, we may say that although there is a slight improvement in the writing skills of the students yet there is no outstanding change in the performance of the students owing to peer assessment and peer feedback. The results suggest that probably in our context these assessment practices may need to be implemented under certain conditions to reap its benefits otherwise it may not prove to be fruitful.

Conclusion

The empirical evidence of this study proved that peer assessment and peer feedback in this particular Pakistani context, may have failed to demonstrate any outstanding learning achievement. Although, the students showed a subtle improvement in their writing skills, but their learning was noticeably insignificant. The performance of the experimental group's students was almost parallel to the performance of the control group's students, who were evaluated by the teacher. It implies that peer assessment and feedback practices did not exceedingly enhance the writing skills of the learners maybe because they rely more on their teacher's assessment and feedback.

Peer assessment done formatively, helps the students to assist one another in planning their learning, identifying their strengths and weaknesses, target areas for remedial action, and develop metacognitive and other personal and professional skills (Topping, 2017) and this is what Pakistani students need to do to excel in their writing skills. It becomes challenging for an individual teacher to provide extensive feedback to a large strength of students in class (Wu & Schunn, 2021), hence, if the students are trained to peer assess one another, they would be more productive through collaborating with one another

and may be able to contribute to one another's learning of writing skills. Hence, peer assessment and peer feedback practices may prove rewarding in our context if they are implemented after rigorous training of the faculty and students. Through this we might be able to reap benefits of peer assessment and feedback to enhance the poor writing skills of our students the way other second language learners across the globe have done.

Recommendations

This study concludes that peer assessment and peer feedback could not largely improve the various aspects of English language writing skills of the students. One of the reasons may be because the students consider teacher's feedback more reliable than their peers' feedback. Hence, in our context of teacher dependent learning environment, a radical shift in the assessment process may prove to be counterproductive. This is supported by Vanderhoven et. al. (2015) who suggests that students used to a teacher-led assessment can get frustrated when the teacher feedback opportunities are totally replaced by peer feedback. Therefore, a mixed model of assessment may be developed where the students may receive scores and feedback from both teacher and peers. The teacher should also provide feedback in addition to peer feedback, as it will provide satisfaction to the students. This corroborates with the ideas of Tsui and Ng (2000) who explain that teacher comments tend to induce more revisions to the macro-structures of a text, whereas peer comments have the specific roles of enhancing a sense of "real" audience in the students, raising the students' awareness of strengths and weaknesses of their own writings, encouraging collaborative learning and fostering an ownership of text.

Our classrooms have a large strength of students, and it becomes difficult for the English language teachers to give individual feedback on the essays (Arshad, 2017). Hence, if we may implement peer assessment and feedback along with the teacher's assessment in English classrooms, it would not only share the teacher's burden but develop a sense of independent learning in the students. However, considering the reliability issues of peer assessment, the students' assessment scores should not be included in the summative assessment. The purpose of implementing peer assessment and peer feedback will be to develop critical ability and taking responsibility of their own learning, in the students. Therefore, peer assessment and feedback may prove useful, if it is introduced as a complementary assessment process, along with the teacher's assessment.

Moreover, the students seemed dissatisfied with their peers' feedback as they felt that the feedback they received, was insufficient. Keeping this in view, it may be suggested that instead of having only one peer assessor, if multiple peer assessors give feedback, then maybe the variety of feedback is more helpful for the student. This is in line with Ashenafi (2017) who explains that students receiving feedback from multiple peers perform complex revisions of their work and produce higher quality products. If one assessor is unable to give satisfactory feedback, then the students may gain a better insight about their errors from the feedback of the other peers.

Another suggestion is intensive training sessions for the students, before commencing peer assessment and peer feedback in the classroom. This may allow the students to learn how to conduct peer assessments and peer feedback sessions. In addition, peer assessment may prove more useful if the students collaborate with the teachers in developing the assessment criteria. These practices involving students as active partners in the assessment process give them the opportunity to develop their capacity to self-regulate their learning (Ion, Sánchez & Agud, 2019).

Considering the results of previous studies based on peer assessment in enhancing the writing skills of the students, it may be concluded that if implemented with proper training with the sole purpose of learning of writing skills, peer assessment and feedback practices may also prove beneficial in Pakistani context. We need to promote these assessment practices as they reduce teacher dependence, helps in sharing the teacher's burden, encourage collaborative learning and develops higher order thinking in the students.

References

- Akhtar, H., Saeed, M., & Ayub, S. (2019). Peer Tutoring: An effective technique to enhance students' English writing skills. *Global Social Sciences Review, IV*, 299-305.
- Akram, M. (2017). Learning and teaching English in Pakistan: predicaments and solutions. *International Journal of Educational Sciences, 19*(1), 10-14.
- Arshad, H. M. (2017). *An exploratory study of primary school teachers' perceptions of group work as a way of teaching English in public sector primary schools of Punjab* (Doctoral dissertation, University of Glasgow).
- Ashenafi, M. M. (2017). Peer-assessment in higher education—twenty-first century practices, challenges and the way forward. *Assessment & Evaluation in Higher Education, 42*(2), 226-251.
- Bordens, K. S., & Abbott, B. B. (2018). *Research design and methods: A process approach (10th ed)*. Mountain View, Calif: Mayfield Pub.
- Bryan, C., & Clegg, K. (Eds.). (2019). *Innovative assessment in Higher Education: A handbook for academic practitioners*. London: Routledge. <https://doi.org/10.4324/9780429506857>
- De Guerrero, M. C., & Villamil, O. S. (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *The Modern Language Journal, 84*(1), 51-68.
- Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review, 32*(2), 481-509.

- Fareed, M., Jamal, U. B., & Zai, R. A. Y. (2021). Peer feedback on writing skills: perceptions of Pakistani ESL postgraduate students. *Pakistan Journal of Educational Research*, 4(1), 399-415.
- Fraenkel, J. R. & Wallen, N.E (2009). *How to Design and Evaluate Research in Education* (7th ed). New York. McGraw-hill.
- Gay, L. R., Mills, G. E., & Airasian, P. W. (2012). *Educational research: Competencies for analysis and applications*. Boston: Pearson.
- Gencha, M. G. (2018). The effect of peer assessment on English language writing instruction & the perceptions of students: the case of selected accounting CEP students at Hawassa University in focus. *IOSR-JHSS Vol*, 23.
- Graneheim, U. H., Lindgren, B. M., & Lundman, B. (2017). Methodological challenges in qualitative content analysis: A discussion paper. *Nurse Education Today*, 56, 29-34.
- Haider, G. (2012). Teaching of writing in Pakistan: A review of major pedagogical trends and issues in teaching of writing. *Journal of Educational and Social Research*, 2(3), 215-215.
- Hill, L. (2016). *Student perceptions of the impact of student-generated feedback in formative essay writing in the english classroom*. [Master's thesis, The University of Waikato]. <http://researchcommons.waikato.ac.nz/>
- Huisman, B., Saab, N., van den Broek, P., & van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: a Meta-Analysis. *Assessment & Evaluation in Higher Education*, 44(6), 863-880.
- Ion, G., Sánchez Martí, A., & Agud Morell, I. (2019). Giving or receiving feedback: which is more beneficial to students' learning? *Assessment & Evaluation in Higher Education*, 44(1), 124-138.
- Jokhio, A. A., Raza, S. S., Younus, M., & Soomro, A. H. (2020). Teaching writing skills in university large classes in pakistan: issues, challenges and solutions. *JEELS (Journal of English Education and Linguistics Studies)*, 7(1), 25-47.
- Khan, H. I. (2011). Testing creative writing in Pakistan: Tensions and potential in classroom practice. *Journal of Humanities and Social Science*, 1(15), 111-119.
- Khurram, B. A. (2018). Promoting learner engagement in a large university-Level ESL class in Pakistan. In *International Perspectives on Teaching English in Difficult Circumstances* (pp. 73-88). Palgrave Macmillan, London.

- Kuyyogsuy, S. (2019). Promoting Peer Feedback in Developing Students' English Writing Ability in L2 Writing Class. *International Education Studies*, 12(9), 76-90.
- 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.
- Marczyk, G., DeMatteo, D., & Festinger, D. (2005). *Essentials of Research Design and Methodology*, John Wiley and Sons. Inc., Hobokon, New Jersey.
- Mayring, P. (2000). Qualitative Content Analysis. *Qualitative Methods in Various Disciplines*, 1(2). DOI <https://doi.org/10.17169/fqs-1.2.1089>.
- Meletiadiou, E. (2021a). Exploring the Impact of Peer Assessment on EFL Students' Writing Performance. *IAFOR Journal of Education*, 9(3), 77-95.
- Meletiadiou, E. (2021b). Opening Pandora's box: how does peer assessment affect EFL students' writing quality? *Languages*, 6(115), 1-17.
- Quynh, N. N. P. (2021, January). Using Peer Assessment in Writing for EFL Learners. In *17th International Conference of the Asia Association of Computer-Assisted Language Learning (Asia CALL 2021)* (pp. 297-302). Atlantis Press.
- Rahman, T. (2010). *Language Policy, Identity and Religion: Aspects of the Civilisation of the Muslims of Pakistan and North India*. Islamabad, Pakistan: Quaid-i-Azam University.
- Rouhi, A., & Azizian, E. (2013). Peer review: Is giving corrective feedback better than receiving it in L2 writing? *Procedia-Social and Behavioral Sciences*, 93, 1349-1354.
- Sarfraz, S., Mansoor, Z., & Tariq, R. (2015). Teachers' and students' perceptions of the communicative language teaching methodology in the CALL environment: A case study. *Procedia-Social and Behavioral Sciences*, 199, 730-736.
- Siddique, M., & Singh, M. K. S. (2017). Role of cooperative learning in enhancing students' writing skills in Pakistani colleges: a review of literature. *Proceedings of the ICECRS*, 1(1).
- Siddiqui, S. (2007). *Rethinking Education in Pakistan: Perceptions, Practices, and Possibilities*. Lahore: Paramount Publishing Press.
- Sultana, M., & Zaki, S. (2015). Proposing project based learning as an alternative to traditional ELT pedagogy at public colleges in Pakistan. *International Journal for Lesson and Learning Studies*, 4(2), 155-173.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of educational Research*, 68(3), 249-276.

- Topping, K. (2017). Peer assessment: Learning by judging and discussing the work of other learners. *Interdisciplinary Education and Psychology*, 1(1), 1-17.
- Topping, K. (2021). Peer Assessment: Channels of Operation. *Education Sciences*, 11(3), 91-105.
- Topping, K. J. (2009). Peer assessment. *Theory into Practice*, 48(1), 20-27.
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9, 147-170.
- Uymaz, E. (2019). The effects of peer feedback on the essay writing performances of EFL students. *International journal of curriculum and instruction*, 11(2), 20-37.
- Vanderhoven, E., Raes, A., Montrieux, H., Rotsaert, T., & Schellens, T. (2015). What if pupils can assess their peers anonymously? A quasi-experimental study. *Computers & Education*, 81, 123-132.
- Vygotsky, L. S. (1978). *Mind in society: development of higher psychological processes* Harvard University Press. Cambridge, Mass.
- Walker, M. (2015). The quality of written peer feedback on undergraduates' draft answers to an assignment, and the use made of the feedback. *Assessment & Evaluation in Higher Education*, 40(2), 232-247.
- Wang, W. (2014). Students' perceptions of rubric-referenced peer feedback on EFL writing: A longitudinal inquiry. *Assessing Writing*, 19, 80-96.
- Wu, Y., & Schunn, C. D. (2021). The effects of providing and receiving peer feedback on writing performance and learning of secondary school students. *American Educational Research Journal*, 58(3), 492-526.
- Yasmin, M., Naseem, F., & Abas, N. (2019). Constraints to developing learner autonomy in Pakistan: university lecturers' perspectives. *Educational Research for Policy and Practice*, 1-18.
- Yüce, E., & AksuAtaç, B. (2019). Peer editing as a way of developing ELT students' writing skills: An action research. *Journal of Language and Linguistic Studies*, 15(4), 1226-1235. Doi: 10.17263/jlls.668377
- Zhang, X. (2018). *An examination of the effectiveness of peer feedback on Chinese university students' English writing performance* (Doctoral dissertation, Oakland University).
- Zheng, L., Cui, P., Li, X., & Huang, R. (2018). Synchronous discussion between assessors and assesseees in web-based peer assessment: impact on writing performance, feedback quality, meta-cognitive awareness and self-efficacy. *Assessment & Evaluation in Higher Education*, 43(3), 500-514.
- Zhu, Q., & Carless, D. (2018). Dialogue within peer feedback processes: Clarification and negotiation of meaning. *Higher Education Research & Development*, 37(4), 883-897.

APPENDIX A

Essay Scoring Rubric

Category	Poor (1)	Needs Improvement (2)	Satisfactory (3)	Good (4)	Excellent (5)	Score
Content	No thesis statement; Not relevant; Not enough to evaluate	Not clear thesis statement; Inadequate development	Clear thesis statement; Limited development of thesis; Mostly relevant to topic; Lacks detail	Clear thesis statement; Thorough development of thesis; No irrelevant sentence; Appropriate length	Good development of thesis statement; completely relevant; suitable length	
Organization / unity	No organization evident; ideas random, related to each other but not totask; no paragraphing; no unity	Some organization; relationship between ideas not evident; no paragraphing/ grouping; suggestion of unity of ideas	Possible attempted introduction, body, conclusion; ideas grouped appropriately; some unity	Suitable paragraphing and topic sentences; basically unified ideas ; follows standard organizational patterns	Highly effective organizational pattern for convincing, persuasive essay; content relevant and effective	
Sequence/ Coherence & Cohesion	Not coherent; no relationship of ideas evident	Partially coherent; limited use of transitions; relationship within and between ideas unclear; may occasionally use appropriate conjunctions	Partially coherent; shows attempt to relate ideas, still ineffective at times; some effective use of logical connectors within paragraphs	Mostly coherent and persuasive, progression of ideas; successful attempts to use logical connectors	Coherent and convincing; good progression of ideas; uses transitional devices and logical connectors	
Vocabulary	Meaning obliterated; extremely limited range; little to no knowledge of appropriate word use regarding meaning	Meaning inhibited; limited range; some patterns of errors may be evident; much repetition	Meaning seldom inhibited; adequate range, variety; appropriately academic; some use of idiomatic expressions	Meaning not inhibited; adequate range, variety; basically idiomatic; infrequent errors in usage	Meaning clear; sophisticated range, variety; often idiomatic; often original, appropriate choices	
Grammar	Dominated by errors; Does not communicate ideas	Problems in simple/complex structures; Frequent errors of agreement, tense and word order.	Effective but simple structures; Some errors of agreement, tense and word order; Meaning clear	Effective complex structures; Few errors of agreement, tense and word order; Meaningful	Good use of complex structure; No errors; Completely meaningful	
Spelling & Punctuation	Little or no command of spelling, punctuation, paragraphing, capitalization	Evidence of developing command of basic mechanical features; frequent, unsystematic errors	Basic punctuation, simple spelling, capitalization, formatting under control; systematic errors	Occasional mistakes in basic mechanics; sophisticated punctuation; few spelling errors	Uses mechanical devices for stylistic purposes; may be error-free	

Total Score: _____