Exploring the Impact of Peer Assessment on EFL Students' Writing Performance

Eleni Meletiadou
London South Bank University, UK

Abstract

Lately, researchers have expressed their concern for EFL students' poor writing performance and exam failure. They have indicated that peer assessment (PA) can be successfully employed to support a better integration of teaching/instruction with assessment of progress in learning. Bearing this in mind, the current study employed a pre-test post-test quasi-experimental design and aimed to explore the effect of PA on 200 Greek Cypriot EFL students' writing performance. These adolescent learners attended two writing classes per week (90 minutes) for a full school year. Teachers received training in PA skills and then had to train their own students. Students were asked to use a PA rubric which was also devised by the researcher but negotiated between the students and their teachers during the training sessions. Paired T-tests were performed to investigate whether students in the control (n=100 students and 10 teachers) and the experimental groups (n=100 students and 10 teachers) enhanced their writing performance comparing their pre- to post-test scores. The study outcomes indicated that PA could have a moderately positive impact on students' writing performance. The use of PA improved students' writing performance in 5 aspects: mechanics, organisation, content, focus, and vocabulary and language use. In response to the need for more experimentation, this study provides recommendations for PA implementation in secondary school EFL writing classes which enable teachers to improve students' writing performance.

Keywords: peer assessment, writing performance, secondary education, EFL learning

Peer assessment (PA), as an alternative form of "assessment for learning" which promotes learner-centered assessment, has drawn considerable attention for more than four decades (Chang & Lin, 2019; Hoffman, 2019; Meletiadou & Tsagari, 2016; Topping, 2018). It is an educational arrangement where students judge a peer's performance quantitatively, for example, by providing a peer with scores or grades, and/or qualitatively, for example, by providing a peer with written or oral feedback (Topping, 2017). PA has significant pedagogical value because it enables learners to take part in assessment by evaluating their peers' learning process and products (Bryan & Clegg, 2019).

According to the literature, PA supports the learning process by providing an intermediate check of student performance against the criteria, accompanied by feedback on strengths, weaknesses, and/or tips for improvement (Panadero, 2016; Topping, 2017). There can also be learning benefits for peer assessors since they are exposed to other ideas and writing samples and are able to internalise the assessment criteria and standards (Smyth & Carless, 2020). However, not all types of peer feedback may lead to an improvement in performance. Researchers describe several conditions under which peer feedback may have a positive influence on learning (Schünemann et al., 2017). Involving learners in the assessment procedure is widely acknowledged as vital to effective self-regulation since it allows learners to identify mistakes and develop strategies to address them (Zamora et al., 2018). However, the development of PA skills is challenging. It requires continuous and repeated practice for learners to become competent peer assessors (Andrade, 2016). Engagement in PA presupposes that teachers can inspire learners and involve them in carefully designed tasks (Race, 2019). Therefore, participation in PA ultimately intends to have a positive influence on their cognitive development and motivation towards learning (Adachi et al., 2018).

This study aims to explore whether PA of writing can be used to improve adolescent students' writing performance and present a PA implementation study in the Greek-Cypriot context since action research in secondary education is scarce. Initially, research studies highlighting the use of PA of writing in EFL contexts will be presented to examine some of the findings in the literature. Then, the methodology of the study will be described, and its findings will be rigorously discussed. Finally, conclusions will be drawn, the limitations of the study will be highlighted and suggestions for further research as well as recommendations will be provided.

Literature Review

To date, literature that empirically links quality criteria for feedback to performance improvement in the case of PA is scarce and few studies adopt a quasi-experimental approach to explore the impact of instructional interventions on PA efficacy and student learning (Double et al., 2020). The present study intended to explore whether the effectiveness of PA as a learning tool could be raised through an innovative instructional intervention in secondary education. It also aimed at developing a deeper learning experience enabling learners to engage with new information in terms of the written assignment, assessment criteria, and the assessment procedure as opposed to repetitive and ultimately unsuccessful learning (Topping, 2017).

There are several principles for effective formative PA which is intended to enhance learning. First, formative PA should aim at self-regulation enabling learners to monitor their learning, set goals, develop suitable strategies, manage resources, and work consistently to achieve these goals. Students should be allowed to take responsibility for and take control of their learning (Lee & Hannafin, 2016). Research indicates that teachers should provide information about

expectations and aims. Further, assessment criteria should be explained clearly. Teachers should cooperate with learners to design assessment rubrics and offer opportunities for learners to provide feedback (marks and comments) to each other in relation to the defined assessment criteria. They should also use PA with their students because it encourages low-achieving learners to work hard and overcome obstacles when learning to write. Learners feel that their peers, who take part in PA, share their concerns and provide them with continuous support (Barrot, 2016). This study intends to offer recommendations for PA implementation which may help teachers use PA of writing effectively in their classes.

One of the goals of using PA in EFL classes is that it can guide learners to reflect more carefully on the same elements of their own written work (Hicks et al., 2016). As learners enhance their writing performance, while comparing their writings and receiving feedback from their friends and teachers, the possibility of engaging in fruitful conversations in terms of which they exchange ideas, clarify points, ask questions, and examine as well as reflect on their options increases (Zhu & Carless, 2018). Peer assessment may foster enhanced learning because students can provide additional feedback. This type of response is different and is possibly received and understood more effectively than teacher feedback (Rotsaert, Panadero & Schellens, 2018). Revisions initiated by teacher feedback were often found to be less successful than those related to peer comments although peer feedback sometimes induced uncertainty (Allen & Mills, 2016). This study indicates ways in which PA can be used by adolescent students who wish to improve their writing skills and become more autonomous learners.

Teacher feedback, although highly evaluated by learners is often associated with confusion, misinterpretation and miscommunication (Edwards & Liu, 2018). On the contrary, peer advice generates discussion and increases reflection as peers ask for clarification and negotiate meaning (Kuyyogsuy, 2019). Peer assessment enhances learners' understanding in terms of writing, allows for more self-corrections, checking books, and asking teachers for clarifications as students are encouraged to assume responsibility for their own assignments (Fan & Xu, 2020). It increases mindful reception as well as the frequency, extent and speed of marks and comments for learners while reducing teachers' workload (Ashenafi, 2017). The current study wishes to explore how involving learners in the assessment procedure may increase the amount and number of assessment opportunities and improve their writing performance. Therefore, the potential lower quality of student feedback may be an acceptable trade-off if PA enhances learner engagement and progress.

Conversely, some studies indicate that learners think that PA is aimless because peers are not regarded as experts, tend to provide positive feedback to friends, and teachers make the final decisions anyway (Wu & Schunn, 2020). It is crucial to remember, however, that formative assessment methods can enrich learners' subsequent performance in summative tests (Dixson & Worrell, 2016). Student engagement in assessment also aims to prepare learners for lifelong learning (Nguyen & Walker, 2016). Consequently, this study seeks to explore ways in which secondary school students can develop their professional skills (i.e., reflection) which are valued by employers.

Peer assessment may also be considered as a luxury or a practice which is, in a way, irrelevant when the aim is enhanced performance in external high-stakes tests (James, et al., 2017). Since peers are not experts, the accuracy of PA varies (Reinholz, 2016). Further, peer assessors' judgement and comments are often challenged by peer assessees as learners do not have the kind of authority and subject knowledge that teachers have (Topping, 2017). However, PA in

this study is used in combination with teacher assessment (TA) as the objective is to complement TA and gradually allow students to develop their writing skills.

To sum up, findings in the literature are quite confusing. Although PA may yield various benefits in relation to student writing performance (Chien et al., 2020), there still seems to be an emphasis on teacher-centred instruction and assessment despite students' poor performance in formal tests of writing in Cyprus and other countries (Tsagari & Meletiadou, 2015). The present study aimed at addressing various literature gaps using a semi-experimental design, rarely used by researchers in the field of PA. It also examined the use of PA in secondary education, a topic which has not been widely explored (Fu et al., 2019). There is still little research on how to adapt this approach to the school contexts of many countries (Topping, 2018).

Therefore, the current study investigated the use of PA as an innovative learning tool which may enhance EFL students' writing skills in secondary education. It also provided PA implementation guidelines for secondary school writing classes. The aim was to enable teachers to improve student performance, particularly in the field of writing. In the present study, the researcher investigated the following research question:

1. What kind of an impact does PA have on adolescent EFL students' writing performance?

Methodology

Participants
The participants in the study were 200, 13-15-year-old students of four public secondary schools in Cyprus. The participating learners faced considerable problems with their writing performance and scored relatively low at the end-of-year exams which gradually prepared them for the International General Certificate of Secondary Education (IGCSE) exams. The learners randomly formed 20 mixed ability EFL groups which attended two 90-minute writing classes per week. This was a convenient sample since the researcher had to work with volunteers, depending on the students and teachers who wished to take part in the study (Mertens & Caskey, 2018).

Participants were all native Greek Cypriots and shared the same cultural and a similar socio-economic background. These students also had a similar kind of exposure to EFL which classified them as intermediate stage (B1) according to the Common European Framework of Reference (CEFR) (Council of Europe, 2001). Participation in the present study was voluntary and conditional on participants and their guardians signing an informed consent form, which had been previously approved by the Cypriot Ministry of Education.

Instrument
This study employed a multiple-trait approach (Nodoushan, 2014) to assessing student essay quality prior to and after receiving training in PA. As for measuring the writing scores of the first drafts and final versions, two different scoring methods were employed: holistic and analytic scoring (Han & Huang, 2017). These two measurements were complementary and provided sufficient information about the participants' writing abilities. The PA rubric (Table 1) was adapted from two lists in White and McGovern (1994) and Jacobs ESL Composition Profile (Jacobs et al., 1981) to reflect learners' errors and their examiners' preoccupations.

Table 1
The PA Rubric

Criteria/Weighting	Excellent/ Very Good	Good	Average	Poor	Very Poor

A. Content (sample statements for all categories)
      1. The main ideas are clear.
B. Organization
      1. The writer uses simple linking devices.
C. Vocabulary and language use
      1. There are subject/verb agreement errors.
D. Mechanics
      1. There are punctuation errors.
E. Focus
      1. There is a clear sense of audience.

Analytic score:
Content: __ (out of 4)
Organization: __ (out of 4)
Vocabulary and language use: __ (out of 4)
Mechanics: __ (out of 4)
Focus: __ (out of 4)
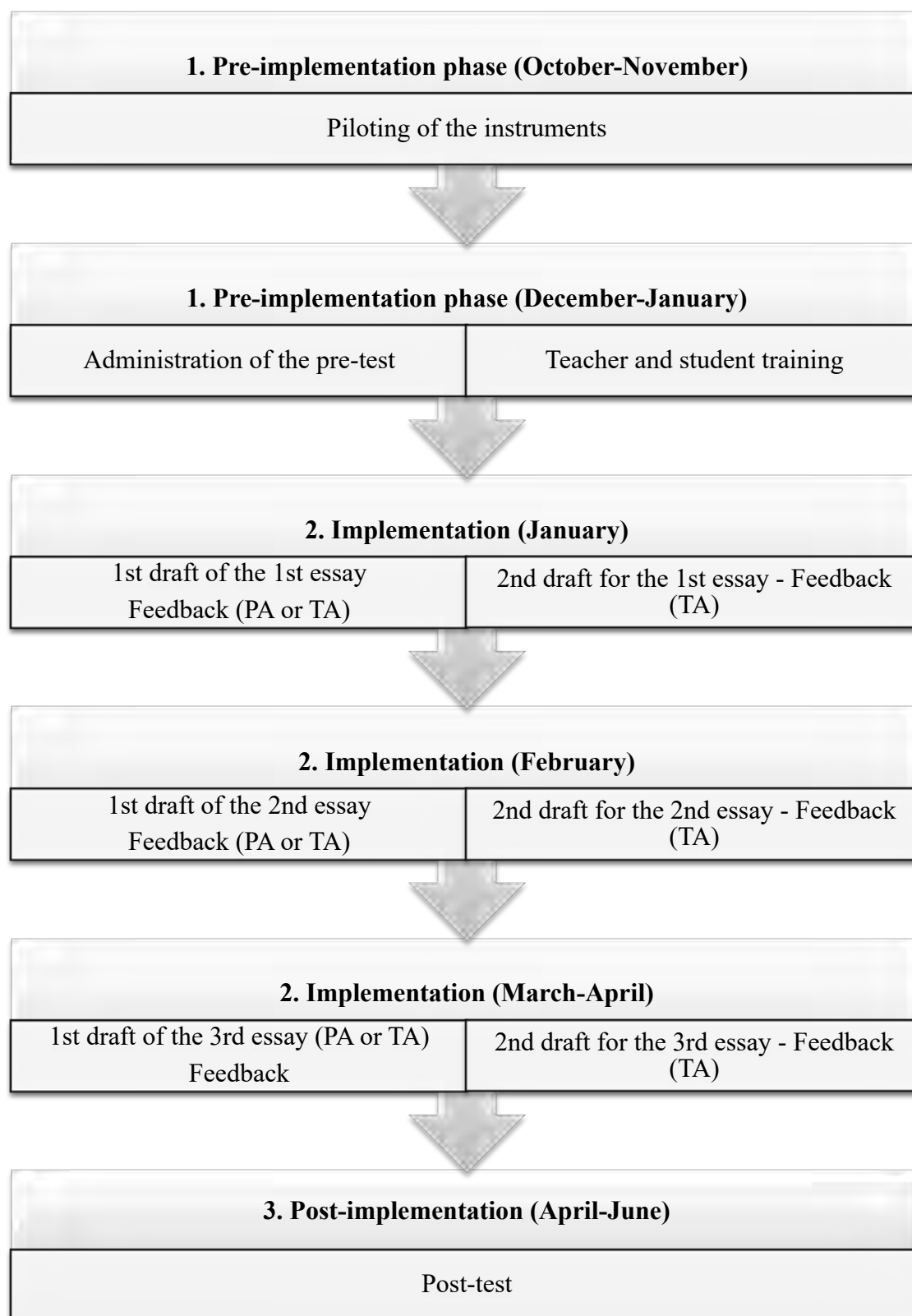**Total score:** __ (out of 20)

Holistic score:

Excellent/Very Good	Good	Average	Poor	Very Poor

Students had to read their peers' essays and reflect on the statements for each category-content, organisation, focus, mechanics, vocabulary, and language use. They had to assign marks for each category (ranging from 0 to 4 marks) and then add them to get a final grade (0-20). Learners were instructed to read the essay under consideration and identify three strengths, three areas for improvement and provide three suggestions for revision. The aim was to introduce learners to PA and help them understand how they could reflect on their peers' and their own essays and offer suggestions to improve their work by carefully editing and proofreading it. The aim was to help students assume a more active role by providing meta-cognitive comments which would allow their peers to revise their work accordingly. The validity of the PA rubric was explored by consultation with experts, including 8 headteachers, one inspector, and 10 qualified EFL teachers who had taught at this level for at least 6 years.

Procedure
The study was divided into three phases (Figure 1). In the pre-implementation phase, all instruments were piloted, and a diagnostic pre-test was administered to ensure that all students who participated in the study were at the intermediate level. Students were provided with a mark but no comments or peer feedback. All essays were marked by their class teachers and an external assessor after a rater calibration session and appropriate training, and 20% of the essays were marked by the researcher.

Figure 1
The Procedure of the Study

1. Pre-implementation phase (October-November)	
Piloting of the instruments	

⬇

1. Pre-implementation phase (December-January)	
Administration of the pre-test	Teacher and student training

⬇

2. Implementation (January)	
1st draft of the 1st essay Feedback (PA or TA)	2nd draft for the 1st essay - Feedback (TA)

⬇

2. Implementation (February)	
1st draft of the 2nd essay Feedback (PA or TA)	2nd draft for the 2nd essay - Feedback (TA)

⬇

2. Implementation (March-April)	
1st draft of the 3rd essay (PA or TA) Feedback	2nd draft for the 3rd essay - Feedback (TA)

⬇

3. Post-implementation (April-June)	
Post-test	

The same test was administered at the end of the study. Students had to write the same type of essay on a different but very similar topic. The aim was to evaluate whether students made any progress after one school year and which groups made more progress, if any, the control or the experimental.

Students took part in writing workshops prior to engaging in PA of their writing assignments. During these workshops, they were instructed how to use the rubric, used it with three sample papers, and received feedback from the instructor on how their comments and marks aligned with the instructor grades and comments on the same essays. These hopefully contributed to a level of comfort and proficiency with the rubric and assignment criteria and possibly enhanced their ability to provide accurate and reliable PAs. Teachers also received extensive training and had to attend weekly sessions because the researcher wanted to check the progress of the implementation. Training and clarity hopefully promoted a trusting environment in the classroom, which had conferred positive outcomes with PA in the literature so far (Hoffman, 2019).

Learners wrote five essays, including a pre-test, 3 essays in two drafts and a post-test. Students in the experimental groups received peer feedback and teacher feedback once each while students in the control groups received teacher feedback twice (Figure 1). All student groups were engaged in the experiment once a week for two teaching sessions (45 minutes each) which added up to approximately 50 teaching sessions. Five compositions (two informal letters as pre- and post-tests, a narrative essay, a descriptive essay, and an argumentative essay) were written in class without disrupting the regular programme to exclude variables such as the amount of time spent on task at home and help from others.

Students wrote the three types of essays after being introduced to the specific genre. Experimental group students received peer feedback and a mark based on the PA rubric and had to assess one of their peers' essays. Students in the control groups received teacher feedback (comments and some corrections of major mistakes) and a grade. Students then received some remedial teaching depending on the challenges they faced in their first draft. Additionally, they were asked to write a second draft. Teachers were instructed to support their students during the whole procedure but not to intervene with their writing. They provided corrections, marks, and comments to students' second draft and after some more remedial teaching they taught the next genre. The researcher collected students' drafts immediately after learners completed each step of the procedure to ensure that teachers followed the researcher's instructions regarding the corrections and comments they provided to their students.

Essays were about 4-5 paragraphs (120-150 words each) long. Instructors monitored the students but were not involved in the actual editing of the essays. Teachers assumed the role of a facilitator by explaining any difficult terms or acting as a consultant by offering advice when needed. All teachers were asked to avoid overcorrecting students' work and provide only occasional basic corrections and comments.

During the feedback sessions, the teacher and student/assessors offered feedback to their students/peers which consisted of both marks and comments based on the PA rubric. In more detail, all experimental group learners devoted 20 minutes of their normal teaching sessions using the PA rubric to assess their peers, while the control groups received teacher feedback. Students were assigned with the correction of their peers randomly and changed every time they had to assess a new draft. The identity of the student/assessor and the student/assessee

were not disclosed to avoid conflicts and bitterness. Anonymity and change of student/assessors also ensured the reliability of the assessment process.

Next, students were asked to re-draft their work. Teacher and peer feedback were provided with a view to improving successive drafts and prompting more revision. Moreover, the feedback sessions were structured tightly regarding time to avoid considerable variation among groups and to increase student concentration. The time between drafts (usually two week) was regarded as sufficient for learners to redraft without feeling undue pressure ensuring the reliability of the assessment process. Additionally, students were asked to peer assess only one draft to avoid any resistance from students.

Learners received remedial teaching depending on their errors. Essentially, the teachers were instructed to use selected parts with significant errors from students' essays and encourage them to identify them and indicate solutions. Teachers also asked students to study their grammar book and the handouts they used to self-correct errors in class at home.

Findings and Discussion

The current study explored the way PA influences student writing performance. Paired T-tests were performed to explore whether students in the control (n=100 students) and the experimental groups (n=100 students) improved their writing performance comparing their pre- and post-test scores (George & Mallery, 2016). These revealed that experimental group students improved their writing performance by 3 marks (out of 20) while control group students improved their performance by only .2 marks. This difference was statistically significant (Table 2). It indicated that students who received PA together with TA showed considerable improvement. This finding was confirmed by previous research (Wanner & Palmer, 2018) which indicated that PA significantly improved the quality of learners' end product from draft to final version.

Table 2
Findings from the Paired T-tests of the Experimental versus Control Groups (overall score)

			M	SD	t	p	Cohen's d
Paired T-tests	Control groups	Post-test	11.73	3.01	.42	.674>.0005	
		Pre-test	11.59	3.5			
	Experimental groups	Post-test	11.73	3.01	.42	674 >.0005	
		Pre-test	10.37	3.22			
Independent T-test (post-test score)	Control groups		11.73	3.01	3.9	.000	.55
	Experimental groups		13.38	2.95			

This study also concluded that PA was not only suitable for adult learners (Baker, 2016). Its adolescent participants provided marks for their peers and improved the marks they received for their own essays through their involvement in the practice of PA and due to the insight they

gained into their peers' work. This allowed them to reflect on their own work and eventually improve it. Taking into consideration that learners can only process feedback for which they are developmentally ready, teacher feedback may often be ineffective. Therefore, teachers should tailor their feedback to student needs but since this is practically impossible especially in large classes, this study has demonstrated a way in which feedback can be individualised for each learner by combining PA, which is more student-friendly, and TA, which is regarded as more precise.

An independent t-test was also performed to explore differences between the post-test scores of both experimental and control groups. On average, in the post-test, experimental group students outperformed control group students (Table 2). Both the difference, which was statistically significant, and Cohen's d indicated that there was a moderately positive impact of PA on students' writing performance (George & Mallery, 2016; Cohen et al., 2013).

Paired t-tests were also performed to explore how improvement spread across the five categories included in the PA rubric (content, organisation, mechanics, focus, vocabulary, and language use) (George & Mallery, 2016). Students were assigned an analytic score per category (0-4 marks). The aim was to further explore which aspect(s) of their writing performance experimental group students improved (Table 3).

Table 3
Findings from the Paired T-tests of the Five Categories (Analytic Scores)

			t	p	Cohen's d
Paired T-tests	Control groups	Mechanics	.00	1.000>.0005	
		Organisation	.42	.675>.0005	
		Focus	2.41	.018>.0005	
		Content	(6.16	.000>.0005	
		Vocabulary and language use	-.420	.675>.0005	
	Experimental groups	Mechanics	7.16	.000 <.0005	.74
		Organisation	5.5	.000<.0005	.59
		Focus	6.16	.000<.0005	.86
		Content	7.08	.000<.0005	.6
		Vocabulary and language use	8.03	.000<.0005	.74

Paired t-tests of the pre- versus post-tests scores control group students received revealed that there was no improvement, which was statistically significant, for mechanics, organisation, focus, content, and vocabulary and language use (Table 3). The same paired t-tests were performed for the experimental groups (George & Mallery, 2016). These revealed that there was improvement which was statistically significant, for mechanics and organisation. However, there was no improvement which was statistically significant for focus, content and vocabulary and language use (Table 3).

Students in the experimental groups improved their vocabulary more than any other aspect of writing (Table 3). This indicates that PA may have a positive impact on different aspects of

students' writing performance. Students read their peers' essays and learnt new words which they then used in their own essays. Learners also improved the content of their essays (t=7.08). They possibly located new ideas in their peers' work and added new content to their own essays. Several studies have investigated the revisions made by learners after receiving PA or TA. These reported that PA leads to more meaning-level revisions while TA leads to more surface-level revisions (Rotsaert et al., 2018). However, none of these studies have indicated that students successfully improved almost all aspects of their writing performance.

In addition, other researchers (Choi, 2013) reported that students, when using peer feedback, mostly concentrated on surface level errors, involving grammatical and spelling mistakes, instead of deep or semantic level issues such as content. In the current study, the impact of peer feedback was detected more on deep and semantic level issues rather than surface level issues possibly because students received training prior to the implementation. This study confirmed some findings from previous research (Lee, 2015) suggesting that peer feedback might ultimately lead to more language improvement, because students are possibly more willing to participate in assessment and learn more easily from their peers since they understand peer feedback better than teacher feedback.

Students, who participated in this study, also managed to upgrade the mechanics of their writing (Table 3) supporting previous research (Yaghoubi & Mobin, 2015). Therefore, they confirmed that the use of PA can cognitively impact how students organise their thoughts as they write. Students had the opportunity to reflect on their work, edit and proofread their essays more carefully after providing feedback to their peers. The findings of this semi-experimental study about mechanics contradict previous research indicating that there are no significant gains for EFL students in terms of mechanics when students use PA (Wanner & Palmer, 2018). These clearly indicate that PA can have a positive impact on students' writing by helping them edit and proofread their work more carefully taking into consideration their peers' comments.

Additionally, students enhanced aspects of their essays related to focus (Table 3). PA allowed students to better understand the texts including the schematic structure and linguistic features of the genre. It successfully raised students' awareness of the context, the reader and facilitated the interpretation of the writer's intended meaning since experimental group students in this study conformed more to the conventions of the genre in hand. The deliberate focus on genres, which were included in the PA rubric, helped learners become more aware of the requirements of the different genres and take them into serious consideration when writing their essays. Consequently, they developed their meta-cognitive skills since they were asked to improve their work relying on two kinds of feedback.

However, students were able to refine their language use and organisational skills (Table 3) less than the other aspects of their writing. Although students looked at their work again, they did not manage to improve their use of grammar and their organisational skills as more time and effort is required to improve these aspects of writing. Previous researchers (Edwards & Liu, 2018) reported that a combined focus on both language form and content leads to greater gains than either focus on form or focus on content alone. This was also confirmed by the current study which showed that feedback on both form and content can result in improvement in all aspects of writing.

When comparing students' performance in the post-test, experimental group students seemed to have improved their performance more in terms of vocabulary, language use and focus (Table 3) rather than in other aspects of writing. This finding indicates that students who used

PA enriched their vocabulary and comprehended the requirements of the specific genre used in both the pre-test and post-test (informal letter) even better. Various measures of text improvement have been employed in different studies, that is, some researchers considered improved grammar as a characteristic of enhanced text quality (Liao, 2016). Other researchers also reported that trained students can provide specific and relevant feedback on global features of writing, such as genre, which in turn may result in better quality in their revised drafts (Subaşı, 2014).

Students also improved certain aspects of their writing, for example, mechanics, content, and organisation (Table 3) slightly less than other aspects. Experimental group students used PA for a few months with only three types of essays. Previous studies deemed organisation of information as an important factor in determining text quality as they have shown that PA can improve students' organisational skills in writing (Hwang, Hung & Chen, 2014). Previous research has also indicated that intermediate EFL students improved their texts significantly in organisation, cohesion and vocabulary when using PA of writing from pre-test to post-test (Ebadi & Rahimi, 2017). However, adolescent intermediate EFL students, who were inexperienced in PA, needed more time and exposure to this approach to improve these aspects of their writing performance.

To sum up, the findings of the present study indicated that students who used PA improved their overall writing performance by 3 marks out of 20, a statistically significant finding, while students who received only TA improved their performance by .14. Finally, students who employed PA improved their writing performance by at least half a mark out of 4 for each one of the categories included in the PA rubric, that is content, organisation, vocabulary and language use, mechanics, and focus. These findings revealed that students, who used PA in their writing classes received multiple benefits in all domains included in the EFL essay rubric. Consequently, since PA-related learning seems to provide multiple benefits to adolescent learners in terms of their final summative assessment, it might be a worthwhile exercise. Therefore, teachers should consider engaging their learners in PA during their classes to enhance their writing performance.

Conclusion and Recommendations

Few studies adopt a quasi-experimental approach to study the impact of instructional interventions on PA effectiveness and learning (Saito, 2012). Previous studies relied only on marks to make claims for the potential benefits of PA. Moreover, the absence of a control group has been the main weakness of longitudinal studies on feedback in writing (Bitchener et al., 2012). The current study relied not only on students' overall marks but also on their analytic scores on five important writing aspects, content, organisation, vocabulary and language use, mechanics, and focus to explore whether PA can have a positive impact on students' writing performance. It used both control and experimental groups in a semi-experimental design to explore in what aspects and to what extent PA could have an impact on intermediate adolescent EFL students' writing performance.

For many instructors, the most viable alternative to teacher feedback would be peer feedback. This has become almost as common as teacher feedback in writing classes (Lee et al., 2016). The current study suggests that the introduction of PA in EFL classes from an early age may improve students' writing performance even more as learners can receive extensive training and devote a considerable amount of time, that is several years, to learn how to use PA effectively. It also recommends the use of PA in combination with TA to maximise the benefits

of PA and allow students to be exposed to multiple types of feedback. These will allow learners to reflect on various aspects of their writing.

Peers and teachers tend to focus on different aspects of writing, when asked to provide feedback, leading to potential differences in improvement in students' writing. For example, it has been suggested that teachers may focus on surface-level issues while peers may focus more on meaning-level issues (Baker, 2016). The positive impact that PA can have on all aspects of students' writing should be taken into consideration by EFL teachers who face considerable challenges as they try to help their students improve their writing performance. Such an outcome conforms to what has been reported by previous studies, namely that PA can engage students in making reflections when they assume the role of tutor as well as tutee (Wang et. al., 2017).

However, while instructive, the findings of this study may not be representative enough to allow generalizations, a challenge to be undertaken in future studies. Although positive effects were found, it became apparent that the training could have been much more systematic and of longer duration than was feasible to organise in the available context and time span of the current study. This study focused only on short-term effects of PA training. Therefore, more structured PA training for both teachers and students and critical reflection on assessment might have had a more powerful effect on students' writing performance.

PA needs to be elaborate and frequent as well as focused on learners' performance, their learning needs, and the actions under learners' control rather than on the learners themselves. PA should also be timely so that learners can reflect on it and use it in their work or ask for help if they need any clarifications. Its goal should be to aid with the assignment at hand and allow learners to understand the assessment criteria even better. It also needs to be suitable to learners' conception of learning and previous knowledge as well as attended to and acted upon.

This study also supports research which suggests that students may ignore or misuse teacher commentary when revising drafts and thus profit when they receive more (peer) feedback (Yu, 2019). More importantly, PA is seen by many researchers as a way of giving more control and autonomy to students. It involves them actively in the feedback process as opposed to a passive reliance on teacher's feedback to 'fix' up their writing (Alzaid, 2017). The findings of this study confirmed that PA can improve EFL learners' writing skills by allowing them to assume responsibility for other students' and their own learning which is in line with previous research (Topping, 2017).

These findings contribute to linguistic theory by suggesting that PA is anticipated to make a significant contribution to the field of education if sufficient training and support is provided to all participants and carefully designed tools are used to familiarize learners with the PA process. Moreover, PA should be introduced gradually and used on a regular basis as early as possible, that is even in primary education, and the emphasis should be on the formative use of PA as an innovative learning-oriented tool employed by teachers to enhance students' skills.

The present study indicated that PA can be a promising alternative assessment method for EFL teachers in secondary education. PA may be used to raise a more open assessment culture and empower adolescent students by involving them in assessment. Finally, it assisted in generating an evidence-based argument regarding the quality of PA as a tool for enhancing EFL writing skills (Wanner & Palmer, 2018).

PA reflects the attempt of the education reform initiatives in many countries, i.e., England and Hong Kong, to move from a testing culture to an assessment culture and promote all round education and life-long learning. Within the sociocultural context of countries like Cyprus, where the stress on measurement and accountability has existed for a long time, the successful implementation of PA in the way it is intended needs promotion of conversations about assessment, teaching and learning among all stakeholders, parents, students, teachers, and senior educational management to promote change, advocate reform, advocate assessment literacy and define the new aims and associated roles.

Research continues to characterize teachers' assessment and evaluation practices as largely incongruent with recommended best practice (Tsagari, 2016). Teachers' assessment il-literacy has resulted in inaccurate assessment of students causing them to fail to reach their full potential. In an article published by Deluca et al. (2016), the lack of assessment literacy was presented as professional suicide. Assessment literacy (AL) is seen, therefore, as a sine qua non for today's competent educator (Levi & Inbar-Lourie, 2020). As such, AL must be a pivotal content area for current and future staff development endeavours. This will allow teachers to familiarise themselves and experiment with a variety of promising alternative assessment methods, such as peer assessment, in their effort to help their students improve their writing performance.

References

Adachi, C., Tai, J. H. M., & Dawson, P. (2018). Academics' perceptions of the benefits and challenges of self and peer assessment in Higher Education. *Assessment & Evaluation in Higher Education*, *43*(2), 294 –306. https://doi.org/10.1080/02602938.2017.1339775

Allen, D., & Mills, A. (2016). The impact of second language proficiency in dyadic peer feedback. *Language Teaching Research*, *20*(4), 498–513. https://doi.org/10.1177/1362168814561902

Alzaid, J. M. (2017). The effect of peer assessment on the evaluation process of students. *International Education Studies*, *10*(6), 159–173. https://doi.org/10.5539/ies.v10n6p159

Andrade, M. S. (2016). Effective organisational structures and processes: Addressing issues of change. *New Directions for Higher Education, 173*, 31–42. https://doi.org/10.1002/he.20177

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. https://doi.org/10.1080/02602938.2015.1100711

Baker, K. M. (2016). Peer review as a strategy for improving students' writing process. *Active Learning in Higher Education*, *17*(3), 179–192. https://doi.org/10.1177/1469787416654794

Barrot, J. S. (2016). Using facebook-based e-portfolio in ESL writing classrooms: Impact and challenges. *Language, Culture and Curriculum*, *29*(3), 286–301. https://doi.org/10.1080/07908318.2016.1143481

Bryan, C., & Clegg, K. (Eds.). (2019). *Innovative assessment in Higher Education: A handbook for academic practitioners*. London: Routledge. https://doi.org/10.4324/9780429506857

Chang, C., & Lin, H. C. K. (2019). Effects of a mobile-based peer-assessment approach on enhancing language-learners' oral proficiency. *Innovations in Education and Teaching International*, 1–12. https://doi.org/10.1080/14703297.2019.1612264

Chien, S. Y., Hwang, G. J., & Jong, M. S. Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education*, *146*, 103751. https://doi.org/10.1016/j.compedu.2019.103751

Choi, J. (2013). Does peer feedback affect L2 writers' L2 learning, composition skills, metacognitive knowledge, and L2 writing anxiety? *English Teaching Forum, 68*(3), 187-–213. https://doi.org/10.15858/engtea.68.3.201309.187

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. New York: Routledge. https://doi.org/10.4324/9780203720967

Council of Europe. (2001). *Common european framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Deluca, C., Lapointe-Mcewan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, *28*(3), 251–272. https://doi.org/10.1007/s11092-015-9233-6

Dixson, D. D., & Worrell, F. C. (2016). Formative and summative assessment in the classroom. *Theory into Practice*, *55*(2), 153–159. https://doi.org/10.1080/00405841.2016.1148989

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. https://doi.org/10.1007/s10648-019-09510-3

Ebadi, S., & Rahimi, M. (2017). Exploring the impact of online peer-editing using google docs on EFL learners' academic writing skills: A mixed methods study. *Computer Assisted Language Learning*, *30*(8), 787–815. https://doi.org/10.1080/09588221.2017.1363056

Edwards, J. H., & Liu, J. (2018). *Peer response in second language writing classrooms*. University of Michigan Press.

Fan, Y., & Xu, J. (2020). Exploring student engagement with peer feedback on L2 writing. *Journal of Second Language Writing*, *50*, 100775. https://doi.org/10.1016/j.jslw.2020.100775

Fu, Q. K., Lin, C. J., & Hwang, G. J. (2019). Research trends and applications of technology-supported peer assessment: A review of selected journal publications from 2007 to 2016. *Journal of Computers in Education*, *6*(2), 191–213. https://doi.org/10.1007/s40692-019-00131-x

George, D., & Mallery, P. (2016). *IBM SPSS statistics 23 step by step: A simple guide and reference*. London: Routledge. https://doi.org/10.4324/9781315545899

Han, T., & Huang, J. (2017). Examining the impact of scoring methods on the institutional EFL writing assessment: A Turkish perspective. *PASAA: Journal of Language Teaching and Learning in Thailand*, *53*, 112–147. https://files.eric.ed.gov/fulltext/EJ1153666.pdf

Hicks, C. M., Pandey, V., Fraser, C. A., & Klemmer, S. (2016). Framing feedback: Choosing review environment features that support high quality peer assessment. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 458 –469). https://doi.org/10.1145/2858036.2858195

Hoffman, B. (2019). The influence of peer assessment training on assessment knowledge and reflective writing skill. *Journal of Applied Research in Higher Education*, *11*(4), 863–875. https://doi.org/10.1108/jarhe-01-2019-0004

Hwang, G. J., Hung, C. M., & Chen, N. S. (2014). Improving learning achievements, motivations, and problem-solving skills through a peer assessment-based game development approach. *Educational Technology Research and Development*, *62*(2), 129–145. https://doi.org/10.1007/s11423-013-9320-7

Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfield, V., & Hughey, J. (Eds). (1981). *Testing ESL composition: A practical approach*. Rowley, Mass: Newbury House.

James, S., Lanham, E., Mak-Hau, V., Pan, L., Wilkin, T., & Wood-Bradley, G. (2018). Identifying items for moderation in a peer assessment framework. *Knowledge-Based Systems*, *162*, 211–219. https://doi.org/10.1016/j.knosys.2018.05.032

Kuyyogsuy, S. (2019). Promoting peer feedback in developing students' English writing ability in L2 writing class. *International Education Studies*, *12*(9), 76–90. https://doi.org/10.5539/ies.v12n9p76

Lee, M. K. (2015). Peer feedback in second language writing: Investigating junior secondary students' perspectives on inter-feedback and intra-feedback. *System*, *55*, 1–10. https://doi.org/10.1016/j.system.2015.08.003

Lee, E., & Hannafin, M. J. (2016). A design framework for enhancing engagement in student-centered learning: Own it, learn it, and share it. *Educational Technology Research and Development*, *64*(4), 707–734. https://doi.org/10.1007/s11423-015-9422-5

Lee, I., Mak, P., & Burns, A. (2016). EFL teachers' attempts at feedback innovation in the writing classroom. *Language Teaching Research*, *20*(2), 248–269. https://doi.org/10.1177/1362168815581007

Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: Learning from the teachers. *Language Assessment Quarterly*, *17*(2), 168–182. https://doi.org/10.1080/15434303.2019.1692347

Liao, H. C. (2016). Enhancing the grammatical accuracy of EFL writing by using an AWE-assisted process approach. *System*, *62*, 77–92. https://doi.org/10.1016/j.system.2016.02.007

Mertens, S. B., & Caskey, M. M. (Eds.). (2018). *Literature reviews in support of the middle level education research agenda*. IAP. https://doi.org/10.1080/00940771.2020.1787750

Meletiadou, E., & Tsagari, D. (2016). The washback effect of peer assessment on adolescent EFL learners in Cyprus. In D. Tsagari (Ed.), Classroom-based assessment in L2 contexts. (pp. 138-160). Newcastle upon Tyne: Cambridge Scholars Press.

Nguyen, T. T., & Walker, M. (2016). Sustainable assessment for lifelong learning. *Assessment & Evaluation in Higher Education*, *41*(1), 97–111. https://doi.org/10.1080/02602938.2014.985632

Nodoushan, M. A. S. (2014). Assessing writing: A review of the main trends. *Studies in English Language and Education*, *1*(2), 116–125. https://doi.org/10.24815/siele.v1i2.1831

Panadero, E. (2016). Is it safe? Social, interpersonal, and human effects of peer assessment. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment*, (pp. 247-266), Routledge.

Race, P. (2019). *The lecturer's toolkit: A practical guide to assessment, learning and teaching*. London: Routledge. https://doi.org/10.4324/9780429060205

Reinholz, D. (2016). The assessment cycle: a model for learning through peer assessment. *Assessment & Evaluation in Higher Education*, *41*(2), 301–315. https://doi.org/10.1080/02602938.2015.1008982

Rotsaert, T., Panadero, E., & Schellens, T. (2018). Anonymity as an instructional scaffold in peer assessment: Its effects on peer feedback quality and evolution in students' perceptions about peer assessment skills. *European Journal of Psychology of Education*, *33*(1), 75–99. https://doi.org/10.1007/s10212-017-0339-8

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, *46*(4), 842–854. https://doi.org/10.1002/tesq.67

Schünemann, N., Spörer, N., Völlinger, V. A., & Brunstein, J. C. (2017). Peer feedback mediates the impact of self-regulation procedures on strategy use and reading comprehension in reciprocal teaching groups. *Instructional Science*, *45*(4), 395–415. https://doi.org/10.1007/s11251-017-9409-1

Smyth, P., & Carless, D. (2020). Theorising how teachers manage the use of exemplars: Towards mediated learning from exemplars. *Assessment & Evaluation in Higher Education*, 1–14. https://doi.org/10.1080/02602938.2020.1781785

Subaşı, G. (2014). What are the effects of written peer feedback training on Turkish ELT students' ability to comment on peer writing? *Pinnacle Educational Research and Development, 3*(9), 1–15. https://www.semanticscholar.org/paper/WHAT-ARE-THE-EFFECTS-OF-WRITTEN-PEER-FEEDBACK-ON-TO-Suba%C5%9F%C4%B1/1429d9b7e7b6c95bb320ad23380dda9f1495cc12

Topping, K. J. (2018). *Using peer assessment to inspire reflection and learning*. Routledge.| https://doi.org/10.4324/9781351256889

Topping, K. J. (2017). Peer assessment: Learning by judging and discussing the work of other learners. *Interdisciplinary Education and Psychology*, *1*(1), 1–17. https://doi.org/10.31532/interdiscipeducpsychol.1.1.007

Tsagari, D. (2016). Assessment orientations of state primary EFL teachers in two Mediterranean countries. *Center for Educational Policy Studies Journal*, *6*(1), 9–30. https://www.researchgate.net/publication/303276727_Dina_Tsagari_Assessment_Orientations_of_State_Primary_EFL_Teachers_in_Two_Mediterranean_Countries

Tsagari, D., & Meletiadou, E. (2015). Peer assessment of adolescent learners' writing performance. *Writing and Pedagogy, Special Topic Issue: Writing Assessment, 7*(2), 305–328. https://doi.org/10.1558/wap.v7i2-3.26457

Wang, S. M., Hou, H. T., & Wu, S. Y. (2017). Analyzing the knowledge construction and cognitive patterns of blog-based instructional activities using four frequent interactive strategies (problem solving, peer assessment, role playing and peer tutoring): A preliminary study. *Educational Technology Research and Development*, *65*(2), 301–323. https://doi.org/10.1007/s11423-016-9471-4

Wanner, T., & Palmer, E. (2018). Formative self-and peer assessment for improved student learning: The crucial factors of design, teacher participation and feedback. *Assessment and Evaluation in Higher Education*, *43*(7), 1032–1047. https://doi.org/10.1080/02602938.2018.1427698

Wu, Y., & Schunn, C. D. (2020). When peers agree, do students listen? The central role of feedback quality and feedback frequency in determining uptake of feedback. *Contemporary Educational Psychology*, *62*, 101897. https://doi.org/10.1016/j.cedpsych.2020.101897

Yaghoubi, A., & Mobin, M. (2015). Portfolio assessment, peer assessment and writing skill improvement. *Theory and Practice in Language Studies, 5*(12), 25042511. https://doi.org/10.17507/tpls.0512.10

Yu, S. (2019). Learning from giving peer feedback on postgraduate theses: Voices from Master's students in the Macau EFL context. *Assessing Writing*, *40*, 42–52. https://doi.org/10.1016/j.asw.2019.03.004

Zamora, Á., Suárez, J. M., & Ardura, D. (2018). Error detection and self-assessment as mechanisms to promote self-regulation of learning among secondary education students. *The Journal of Educational Research*, *111*(2), 175–185. https://doi.org/10.1080/00220671.2016.1225657

Zhu, Q., & Carless, D. (2018). Dialogue within peer feedback processes: Clarification and negotiation of meaning. *Higher Education Research & Development*, *37*(4), 883–897. https://doi.org/10.1080/07294360.2018.1446417

Corresponding author: Eleni Meletiadou
Email: elenim@outlook.com