Content list available at http://ijltr.urmia.ac.ir

Iranian Journal of Language Teaching Research

ORIGINAL ARTICLE

Urmia University

Peer Feedback Accuracy in Synchronous and Asynchronous Computer-mediated Conditions in an EFL Context

Saeed Mehrpour [a], Mohammad Hamed Hoomanfard [b, *], Elham Vazin [c]

[a] *Istanbul Aydin University, Turkey*

[b] *Payame Noor University, Iran*

[c] *Shiraz University, Iran*

ABSTRACT

The present study examined peer feedback accuracy in synchronous and asynchronous peer feedback conditions. Employing a counterbalanced repeated-measures design, the researchers examined the accuracy of 2327 comments provided by 96 intermediate EFL (English as a Foreign Language) learners in synchronous and asynchronous computer-mediated conditions. In the synchronous condition, the learners exchanged comments using the desktop version of WhatsApp, and the participants in the asynchronous condition created weblogs and provided feedback in the comment box. All students provided peer feedback in both conditions, and the accuracy of their comments in these computer-mediated environments was compared. Findings showed that the asynchronous condition yielded significantly more accurate comments and fewer missed erroneous items. The data analysis also indicated that low-intermediate students provided more accurate comments in the asynchronous condition. Further, the findings revealed that EFL learners provided comments supported by references to external sources, including textbooks, online sources, teachers' lectures (referentially-justified peer feedback), which were significantly more accurate than those comments not accompanied by references. The findings of this study suggest that when the accuracy of peer feedback is of paramount importance, asynchronous peer feedback design should be implemented.

Keywords: peer feedback accuracy; synchronous; asynchronous; computer-mediated communication; referentially-justified feedback

ARTICLE HISTORY

Received: 29 July 2020 **Revised version received:** 28 June 2021

Accepted: 13 Nov. 2021 **Available online:** 1 Jan. 2023

* Corresponding author: Payame Noor University, Tehran, Iran
Email address: hamedhooman@gmail.com

10.30466/ijltr.2023.121274

Introduction

Although theoretical rationale, empirical evidence, and practical benefits of peer feedback have encouraged instructors to employ this instructional activity in second/foreign language writing programs (Shang, 2017; Tian & Li, 2018), it has been challenged for students' overemphasis on local aspects, negative perceptions of feedback usefulness, and inaccuracy of comments (Kim, 2015; Leki, 1990; Liu & Hansen, 2018; Zhao, 2010). However, due to the plethora of advantages peer feedback brings to writing classes (e.g., enabling learners to engage in the writing process, having a sense of audience, understanding good performance criteria, developing self-assessment ability), researchers have examined different educational conditions to minimize the identified challenges of employing peer feedback in L2 (second/foreign language) writing classes.

A significant factor affecting the quality and acceptability of peer feedback activities is the accuracy of peer comments (Cho & Schunn, 2007; Kaufman & Schunn, 2011; Rotsaert, Panadero, Estrada, & Schellens, 2017). Although the accuracy of peer feedback has long been mentioned as the main drawback of peer feedback (Hovardas, Tsivitanidou, & Zacharia, 2014; Ware & O'Dowd, 2008), it has remained an under-explored area in the peer feedback literature. Previous studies have argued that different factors (e.g., students' learning styles, writing ability, training) can increase the number of accurate comments provided by students; however, to the best of the researchers' knowledge, no prior study has investigated whether and how synchronous and asynchronous computer-mediated modes can result in more accurate peer comments.

The employment of computer-mediated communication technologies to improve language learning conditions is motivated by both theoretical considerations and empirical studies. According to the ecological perspective of language learning, the condition of learning context can, to a large degree, affect the learning process and product (van Lier, 2004). Similarly, the sociocultural theory posits that in the process of moving from other-regulation to self-regulation, mediational tools such as dialogues (Lantolf, 2011; Rastgou, Storch, & Knoch, 2020) and computerized technologies (Gedera, 2016) can play significant roles in learners' cognitive development. Similarly, in the field of feedback, Ellis (2010, p. 338) has signified the role of context as a mediating factor "between the CF [corrective feedback] that learners receive and their engagement with the CF." Ellis argues that the context in which feedback is exchanged can significantly affect learning outcomes. Supporting these theoretical considerations, the findings of the past studies on computer-mediated peer feedback have also shown that the medium through which the communication occurs can affect both the feedback process and product (Cha, 2007; Chang, 2009; 2012; Hoomanfard & Rahimi, 2020; Shang, 2017). On these grounds, examining the frequencies of accurate comments generated in synchronous and asynchronous peer feedback conditions, the present study seeks to illuminate how the use of technology can improve the accuracy of students' peer feedback.

The second part of this study deals with the content of feedback provided in these two conditions. Previous studies have indicated that the content of peer feedback can affect its recipients' perceptions and performance (Hattie & Timperley, 2007; Narciss, 2008). The findings of two studies conducted by Gielen, Peeters, Dochy, Onghena, and Struyven (2010) and Walker (2015) showed that the comments that included arguments, explanations, or reasons provided in support of a specific evaluation significantly improved students' feedback acceptability, perceptions, and revision performance. Informed by these findings and having witnessed EFL students' use of references to reliable sources such as textbooks, online websites, teachers' lectures to make their comments seem more credible (Hoomanfard & Rahimi, 2020; Sánchez-Naranjo, 2019), the researchers have hypothesized that there might be a relationship between students' use of *referentially-justified comments* (the researchers of the present study would like to define it as comments which include elaborated specific guidance necessarily accompanied by a reference to a reliable external source) and the accuracy of peer comments. This hypothesis is also informed by the previous studies showing that while the information provided by a teacher, a book, or a parent is

regarded as credible, peer-generated information is not always perceived as a reliable source (Brown, Irving, Peterson, & Hirschfeld, 2009; Gao, Schunn, & Yu, 2017; Nguyen, 2018; Strijbos, Narciss, & Dünnebier, 2010); thus, references to authoritative sources are likely to make comments seem more credible. In this study, the relationship between students' use of references to credible sources (books, websites, teacher, etc.) in their comments and the accuracy of comments in two synchronous and asynchronous conditions are examined.

The findings of this study can contribute to the theory of peer feedback as it examines peer feedback accuracy, which can have profound effects on students' perceptions of peer feedback accuracy, engagement with feedback, and writing improvement. The research results can provide L2 writing researchers and teachers with insights into how asynchronous computer-mediated communication technology can minimize a major drawback of peer feedback. Furthermore, this study introduces a new feedback type, referentially-justified peer feedback, which can initiate a new area of research on peer feedback in first and second language learning contexts.

Literature Review

Feedback under Synchronous and Asynchronous Conditions

While a large number of studies have investigated computer-mediated feedback on L2 students' texts (see Bahari, 2021 for a comprehensive review), few studies have compared comments on L2 students' written texts provided under synchronous and asynchronous conditions. Lin (2005), examining synchronous and asynchronous peer feedback conditions, found that while the synchronous condition helped learners brainstorm more efficiently, it led to more instances of off-task interactions. The data analysis indicated that while different conditions shaped the way students commented in terms of language functions and conversational focuses, the modes of peer feedback did not affect students' subsequent revisions directly. In a Korean context, Cha (2007) compared EFL students' performance in synchronous chatting and bulletin board (the asynchronous tool). The findings of her study showed that significantly more comments were provided in the asynchronous condition, and this condition brought about more comments addressing the content of texts. Chang (2009) compared EFL learners' interactions and comment types provided in synchronous and asynchronous conditions. She found that in the former condition, learners were more involved in off-task interactions, and they provided more local comments than those in the asynchronous group. As an extension to this study, Chang (2012) found that asynchronous mode generated slightly more on-task interactions than the synchronous one. The asynchronous mode showed its superiority in the frequency of revision-oriented comments; however, the number of comments on local aspects was significantly higher in the asynchronous mode.

Shintani (2015) also found that both synchronous and asynchronous peer feedback conditions improved learners' noticing-the-gap, but self-correction was more successful in the synchronous condition. She also found that focus on meaning and form occurred simultaneously in the synchronous condition, but it took place separately in the asynchronous condition. In a more recent study, Shang (2017) found that asynchronous peer feedback was more successful than synchronous peer feedback in improving the syntactic complexity of EFL students' texts. As another piece of evidence for the effect of the learning condition on the process of peer feedback exchange, Pham (2020) found that the asynchronous nature of Google Docs enabled students to spend more time on providing thoughtful and detailed comments.

This brief review of the literature on the studies comparing synchronous and asynchronous peer feedback showed that different conditions could affect the process (interactions, feedback types, revision quality) and product (writing ability) of feedback exchange activities. Although these studies

have widened our understanding of the affordances of peer feedback in synchronous and asynchronous conditions, they have both led to mixed results in some cases (e.g., the revision quality in synchronous and asynchronous conditions) and have left some gaps in the literature unfilled. One of the niches which has not been occupied thus far is the examination of the accuracy of peer comments generated in synchronous and asynchronous feedback conditions.

Peer Feedback Accuracy

The literature on peer feedback accuracy in both L1 and L2 contexts includes studies that have cautioned us against the detrimental effects of inaccurate peer comments. Prior studies have shown that low peer feedback accuracy can have adverse effects on students' perceptions of peer feedback (e.g., Cho & Schunn, 2007; Kaufman & Schunn, 2011; Rotsaert et al., 2017; van der Kleij & Lipnevich, 2020) and recipients' revision quality (Gielen et al., 2010; Sluijsmans, Brand-Gruwel, & van Merriënboer, 2002). Furthermore, the inaccuracy of peer comments can demotivate teachers to employ peer feedback in their classes (Liu & Carless, 2006; Panadero & Brown, 2017).

Some prior studies (Huisman, van den Broek, & van Driel, 2019; Patchan, Hawk, Stevens, & Schunn, 2013; Patchan & Schunn, 2015) have shown that learners perceive comments provided by less competent students to be less accurate. Chong's (2017) study indicated that EFL students' linguistic accuracy significantly correlated with peer feedback accuracy. The complexity of the learning task is another determining factor affecting the accuracy of peer comments (van Zundert, Könings, Sluijsmans, & van Merriënboer, 2012). In a recent study, Alqassab, Strijbos, and Ufer (2019) found that feedback providers' perceptions of their peer feedback message and anxiety predicted their peer feedback accuracy.

Some other scholars have tried to find ways to minimize inaccurate comments. For instance, Rahimi (2013) found the positive effect of feedback provision training on the accuracy of EFL students' peer comments. Panadero, Romero, and Strijbos (2013) found the positive effect of students' use of evaluation rubrics on the accuracy of their comments. In another study, Yeh, Lo, and Chu (2014), employing an online system, found that students' exposure to their teachers' error correction samples scaffolded their peer feedback providing ability and enabled them to provide significantly more accurate comments in comparison to their initial state at the beginning of the study.

Peer Feedback Content

Following the comprehensive study by Narciss (2008), which shed light on the content of feedback as one of the determining factors in the success of peer feedback, some researchers have conducted empirical studies to investigate how the changes in peer feedback content can affect students' performance. One of these studies, conducted by Strijbos et al. (2010), investigated the difference between the effects of elaborated specific and concise general feedback provided by high-level and low-level peers on graduate students' perceptions of peer feedback and revision performance. Their study showed that the students in the concise general feedback condition outperformed those in the elaborated specific group. Regarding the source of feedback, elaborated specific feedback provided by intermediate peers was regarded as the most adequate. Another study was conducted by Berndt, Strijbos, and Fischer (2018), which investigated the effect of peer feedback content (concise general versus elaborated specific feedback) and sender's competence on students' perceptions of peer feedback, revision performance, and mindful cognitive processing. The results showed that feedback provided by a highly competent peer was perceived as more adequate. The students in the elaborated specific feedback group gave higher scores to positive attitudes toward peer feedback. Regarding revision performance, however, no significant difference was traced.

In few studies, the effects of justified peer feedback were investigated. For instance, Gielen et al. (2010) found that receiving justified feedback (i.e., arguments, explanations, or reasons provided in support of a specific evaluation) significantly improved students' revision performance. Likewise, Walker (2015) found that justified comments (explanation of a correction to a content or skills shortcoming, or explanation of why something is praiseworthy) resulted in the highest number of changes and adaptations in texts. Bolzer, Strijbos, and Fischer (2015) also examined the performance of tertiary level students who were assigned to elaborated specific and elaborated specific plus justification feedback groups. They found that students in the former group spent more time reading peer comments and essays and exerted more effort to make a connection between the comments and the text; however, they did not find any significant difference between the revision performance of students in the two groups.

By adding the *references* element to elaborated specific plus justification feedback presented in previous studies (Bolzer et al., 2015; Gielen et al., 2010; Walker, 2015), the researchers of the current study have aimed to extend this line of research. Such being the case, this study examines whether *referentially-justified feedback* is more accurate than those not accompanied by such references. To the best of our knowledge, no previous study, in either L1 or L2 contexts, has investigated the effect of referentially-justified feedback on students' feedback accuracy.

In addition, while the reviewed studies have contributed to our understanding of computer-mediated peer feedback and peer feedback accuracy, no systematic study has examined the accuracy of peer comments provided in synchronous and asynchronous computer-mediated conditions. Furthermore, the relationship between students' use of referentially-justified comments and their accuracy is another novel issue that is addressed in present study. Based on the reviewed theoretical considerations and the niches in the literature, this study is aimed at answering the following research questions.

Research question 1: Is there any significant difference between the effect of using synchronous and asynchronous computer-mediated conditions on EFL students' peer feedback accuracy?

Research question 2: Are *referentially-justified comments* significantly more accurate than *non-referentially-justified comments*?

Method

Design

The present study employed a repeated-measures design to collect the required data. All the 96 participants experienced both synchronous and asynchronous conditions during the study. This research design, which has been employed by several prior studies on feedback (Chang, 2010; Pham, 2020; Shang, 2017; Shintani, 2015; Warschauer, 1996), is usually utilized to eliminate the problem of prior differences in the groups that can confound the findings in between-subjects designs. The researchers of this study used this research design since there are several variables such as computer literacy, writing ability, reading ability, learning style, working memory capacity, and implicit theories of intelligence (Bahari, 2021; Rahimi, 2015) that could affect the process and product of peer feedback activity. As a result, they used the within-subjects design to ensure that the results gained for each condition could not be attributed to preexisting variables. Furthermore, a counterbalancing design was used to remove the order effect of the conditions. As Ary, Jacobs, Irvine, and Walker (2010) state, "when intact classes must be used, counterbalancing provides an opportunity to rotate out any differences that might exist between the groups" (p. 319). Table 1 provides a brief description of the design of this study.

Table 1
Design of the Study

Sessions 4-11	Essay writing:	Group 1: Synchronous feedback Group 2: Asynchronous feedback
Session 11	Data collection 1	
Sessions 12-18	Essay writing:	Group 1: Asynchronous feedback Group 2: Synchronous feedback
Session 18	Data collection 2	

Participants

Ninety-six EFL students (54 females and 42 males), ranging in age from 20 to 31 attending six intermediate-to-advanced writing classes (as a part of a general English program) at a private language institute in Shiraz, participated in this study. These students were selected based on convenience sampling. Intermediate-level students were selected since they were proficient enough to discuss different aspects of writing in English, and there was still room for their writing ability improvement. All participants were native speakers of Persian. To determine the participants' English language ability, the researchers employed an IELTS Mock Test to assess their English language ability. All of the participants were independent English users with scores ranging between 4.5 and 6.5 band scores, which represent level B in the CEFR scale (Council of Europe, 2001). The mean score of these participants was 5.44 (SD= 0.83). Further, since level B in the CEFR scale is broad enough to include L2 language learners with noticeably different linguistic abilities, the researchers categorized them based on their IELTS exam scores into two low and high-intermediate groups. Those who scored between 4.5 and 5.5 (M=5.01, SD= .42) were regarded as low-intermediate students, and those students with scores over 5.5 (M= 6.03, SD= .13) were considered high-intermediate students.

The participants reported having access to computers (laptop and/or desktop computers). The students could choose other classes in which traditional face-to-face peer feedback was practiced, but all of them welcomed the computer-mediated peer feedback option, and there was no participant attrition during the study. Although the technologies employed in this study were user-friendly, one of the researchers explained the details of using the desktop version of *WhatsApp* and weblogs thoroughly in the classroom. Administering brief one-on-one tests, the researchers ensured that all students had sufficient computer literacy to accomplish peer feedback tasks.

Instruments

Writing tasks

The argumentative writing tasks employed to collect students' feedback were the same in both groups. The first task was used in the 11th session and the second one in the 18th session. The participants had 40 minutes to write an essay containing at least 250 words to argue for or against a topic. The tasks are provided in Appendix A.

Synchronous and Asynchronous Media

Synchronous peer feedback was performed by an application, *WhatsApp*, which is easy to install and use on both smartphones and computers. This messenger was utilized to provide the learners with the platform to paste their texts and discuss the comments. Furthermore, as the sending time of each message could be observed, the researchers could make sure that the interactions involved were synchronous and finished in the specified limited amount of time. To compare the findings

with those of the prior studies, the researchers asked the participants to use the desktop version of this messenger.

For the asynchronous peer feedback condition, the language learners were asked to create simple weblogs. The participants were free to choose any weblog service they wished, but the weblogs had to be accessible to fellow students, and the comment section had to be made enabled. The comments and the writers' responses to them had to be kept under each text so that the teacher (and the researchers) could access and examine them.

Classroom Procedures

The same textbook and the same units were covered in all classes. The students took part in 18 class sessions in a three-month term. In the first three sessions, the students practiced an array of paragraph writing styles and received a short period of peer feedback training. From the fourth session, the learners were taught how to write problem and solution, discursive, and cause and effect essays. The second author provided the students with a short peer feedback training course based on the training procedure recommended in Rahimi (2013).

Since preexisting variables (e.g., writing ability, learning style, implicit theories of intelligence) could affect the process and product of feedback activity, the researchers examined the accuracy of comments provided by the same group in both synchronous and asynchronous conditions. To avoid order effect and ensure the internal validity of the study (Corriero, 2018), the researchers employed a counterbalancing research design. Three classes underwent synchronous treatment in the first half of the term, and the other three classes exchanged comments asynchronously. Then, the participants switched to the alternative treatment after the 11th session. In one of the groups (three classes), there were 46 students, and 50 students (three classes) formed the other group. In both groups, the students provided peer feedback 15 times (sessions 4 to 18). In addition, the researchers assiduously did their best to keep the procedures the same to avoid any possible effect of extraneous variables.

In the synchronous condition, the learners had to write their texts in three days. During the fourth day, they were required to go online and paste their texts onto the messenger window. They were asked to provide feedback and discuss the strengths and weaknesses of each text in at most 30 minutes. Finally, the students had to submit the first and the revised versions. In addition, to minimize the off-task interactions, the participants had to provide a print out of their conversations and submit it to the teacher in the following session. The learners in the synchronous condition were asked to provide local feedback in parentheses immediately after the erroneous item or at the end of paragraphs.

In the asynchronous condition, the participants went through almost the same procedure. They had to write their texts within three days, and then they were asked to post them on their weblogs. In this condition, the participants had to provide comments on their peers' texts in the comment section positioned under the text within two days. The writers and reviewers had two days to discuss the comments and revise their papers. They had to submit their texts to the teacher by the beginning of the following session. Table 2 presents a description of peer feedback activities in both conditions.

Table 2
Details of Peer Feedback Exchange Procedures in Synchronous and Asynchronous Conditions

	Writing the text	Peer reviewing and discussing	Revising and submitting
Synchronous	3 days	30 minutes (for each text)	2 days
Asynchronous	3 days	Two days	2 days

Data Collection

The data collection of this study started with administering an IELTS Mock Test in the first week. The students took the listening and reading sections of the IELTS Mock Test in the first session. On the same day, speaking and writing sections were administered online. The data required to answer the research questions were collected in the 11th and 18th sessions. The researchers collected all the texts and comments provided in these two sessions for further analysis.

Data Analysis

The scoring of the students' reading and listening performances in the IELTS exam was straightforward; however, the participants' speaking and writing performances were scored by one of the researchers of this study and an experienced IELTS examiner (out of the research team), and the inter-rater values for the speaking and writing sections were .86 and .91, respectively.

The data collected to answer the research questions underwent different analysis steps. To examine the accuracy of comments in different computer-mediated conditions, the researchers first categorized the comments based on the three main speech functions (expressive, referential, and directive), suggested by Holmes (2001) and used in prior feedback studies (Jafarigohar, Hoomanfard, & Jalilifar, 2018; Kumar & Stracke, 2007; Stracke & Kumar, 2010). The researchers excluded the expressive comments from the corpus since they represented feedback providers' feelings, which could not be evaluated as accurate or inaccurate. A detailed account of this process is provided in Appendix B. Moreover, the accuracy of the directive and referential comments was scrutinized.

To check the accuracy of comments, the data were first analyzed for the accuracy of the provided comments and then analyzed again for the missed errors. In this study, feedback accuracy referred to the extent to which ideas provided by feedback providers in the form of comments were relevant and accurate. Therefore, the comments with irrelevant or inaccurate statements were coded as inaccurate. The formal comments such as "*put a semi-colon after however*", "*the spelling of knowledgable* [sic] *is wrong*", and "*Don't use from after enjoy!*" were easy to be codified into two categories of accurate and inaccurate since there were specific rules against which these comments could be examined. The second and the third researchers of this study examined the comments independently, and a high coefficient of consistency (Cohen's Kappa) was obtained (98%) for the formal comments. However, due to the nature of global aspects of writing (content and organization), the process of coding the comments was more subjective. For instance, a comment like "*You did not support your topic sentence sufficiently. Solve this problem.*" Or "*Your topic sentence is too broad. Please narrow it down*" were coded as accurate or inaccurate based on the researchers' evaluation of the paragraph. Each comment which caused disparity was discussed until a complete agreement was achieved. After that, another researcher (out of the research team) coded comments as either accurate or inaccurate. The inter-coder reliability between this coding and that of the researchers yielded a coefficient of .92. During extensive discussions, the researchers reached unanimous decisions for each item.

Moreover, following the lead of Hovardas et al. (2014) and Yeh et al. (2014), the researchers employed expert evaluation as the benchmark to examine whether any erroneous item (erroneous

item refers to any substandard item pertinent to content, organization, structure, or mechanical issues) was missed by peer feedback providers. In so doing, two PhD holders provided comments on each text separately. The same comments were provided in more than 96% of cases (Cohen's Kappa). After discussing the disagreements and reaching a specific decision for the necessity of these comments, their comments were compared with those of the students to find out the number of missed erroneous items. Considering both accuracy criteria, five persons (three authors and two PhD holders out of the research team) formed two panels and labeled the comments accurate/inaccurate in four gatherings in a week.

Finally, those comments which had mentions of ideas provided by an external source (a book, a teacher, a website, etc.) along with the citation of the source name or address were coded as referentially-justified comments. Those comments that included texts copied from a book without mentioning the source were not coded as referentially-justified since the reference was not stated. For instance, "*As Mr. X said, we need to write a topic sentence for each paragraph*." and "*Based on page 48 of English Grammar in Use, you have to use the base form of the verb after used to*." were coded as referentially-justified.

Findings

Accurate Comments Provided in Synchronous and Asynchronous Conditions

Excluding 309 expressive comments from the corpus, 2327 comments in the form of referential (n=822) and directive (n= 1505) comments were examined. First, the frequencies of accurate comments provided in synchronous and asynchronous peer feedback conditions were compared (Table 3).

Table 3
Chi-square for Accuracy of Peer Feedback in Synchronous and Asynchronous Conditions

Syn.	Asyn.	X^2	Sig.
751/1034 (72.63%)	1127/1293 (87.1%)	77.9	.000

Table 3 indicates that the participants provided 751 (72.63%) accurate comments in the synchronous condition and 1127 (87.1%) accurate comments in the asynchronous condition. In other words, the number of accurate comments provided in the asynchronous condition was significantly higher than that in the synchronous condition ($X^2(1) = 77.9$, $p<.01$).

To have a better understanding of whether and how different computer-mediated peer feedback conditions enabled high and low-intermediate students to provide accurate comments, the proportional accuracy of comments and missed errors were compared.

Table 4
Chi-square for Accuracy of Peer Feedback Provided by High and Low-intermediate Students in Synchronous and Asynchronous Conditions

HL				LL			
Syn.	Asyn.	X^2	Sig.	Syn.	Asyn.	X^2	Sig.
516/673 (76.7%)	693/786 (88.1%)	33.74	.000	235/361 (65.1%)	434/507 (85.6%)	50.17	.000

As shown in Table 4 and Figure 1, high and low-intermediate students provided significantly higher levels of accurate comments in the asynchronous feedback condition, high-intermediate= 88.1% and low-intermediate= 85.6%, than in the synchronous condition, high-intermediate=76.7% and low-intermediate= 65.1%, ($X^2_{high\text{-}intermediate}$= 33.74 and $X^2_{low\text{-}intermediate}$= 50.17, p<.01).

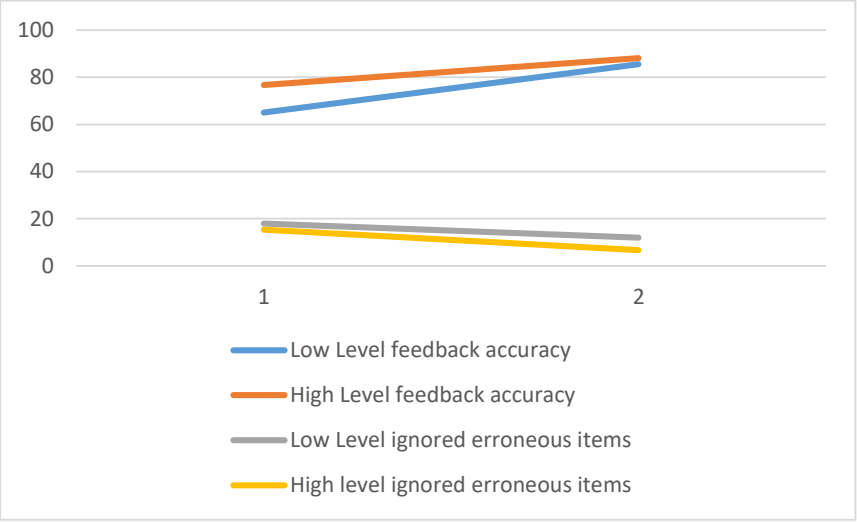

Figure 1. Feedback Accuracy by High and Low-intermediate Students in Synchronous and Asynchronous Conditions and their Ignored Erroneous Items

Table 5 provides information about the missed erroneous items in both synchronous and asynchronous conditions. The Chi-square results showed that both high and low-intermediate students missed fewer erroneous items in the asynchronous condition ($X^2_{high\text{-}intermediate}$= 31.9 and $X^2_{low\text{-}intermediate}$= 7.15, *p*<.01).

Table 5
Chi-square for Missed Errors in Synchronous and Asynchronous Conditions by High and Low-intermediate Students

HL				LL			
Syn.	Asyn.	X^2	Sig.	Syn.	Asyn.	X^2	Sig.
122/795 (15.31%)	56/842 (6.65%)	31.9	.000	79/440 (17.95%)	69/576 (12.0%)	7.15	.007

Accuracy of Referentially-justified Comments

The analysis of the provided comments in the synchronous and asynchronous conditions indicated that around 781 comments (33.56%) were referentially-justified, and 1546 comments (66.44%) were not accompanied by references to external sources (Table 6).

Table 6
The Frequencies and Percentages of Referentially and Non-referentially-justified Comments

All comments (2327)			
Referentially-justified 781 (33.56%)		Non-referentially-justified 1546 (66.44%)	
Accurate	Inaccurate	Accurate	Inaccurate
664 (85%)	117 (15%)	1214 (78.5%)	332 (21.5%)

The analysis of the data indicated that 85 percent (n=664) of referentially-justified comments and 78.5 percent (n=1214) of comments without any reference were accurate, yielding a significant difference between these two sets of comments (X^2= 14.05, p<.001).

In total, 781 comments were accompanied by references, and 236 of them (30.21%) were provided in the synchronous group and 545 (69.78%) in the asynchronous condition. The results of Chi-square tests showed that the participants provided significantly more references in the asynchronous condition than in the synchronous condition (X^2=96.23, p<.001). This result suggested that different feedback exchange environments resulted in different levels of referentially-justified feedback provision. The referentially-justified comments were in three main categories: references to their teacher ($N_{synchronous}$=168, 71.18%-- $N_{asynchronous}$=201, 36.88%), references to their books ($N_{synchronous}$=32, 13.55%--$N_{asynchronous}$=106, 19.44%), and references to online resources ($N_{synchronous}$=36, 15.25%--$N_{asynchronous}$=238, 43.66%).

Discussion

The present study had a two-fold objective. First, the researchers examined the accuracy of peer comments provided by EFL students in synchronous and asynchronous computer-mediated conditions. This research project was motivated by previous studies which had shown that different computer-mediated conditions could affect the process and product of peer feedback activities (Chang, 2009; 2012; Hoomanfard & Rahimi, 2020; Lin, 2005; Pham, 2020; Shang, 2017; Shintani, 2015). The analysis of the data indicated the superiority of the asynchronous computer-mediated condition in enabling intermediate EFL students to provide more accurate comments and miss fewer errors.

Based on the findings of the study, the differential effects of synchronous and asynchronous conditions on learners' peer feedback accuracy are in line with the ecological perspective of language learning and sociocultural theory, which posit that learners' cognition and performance are significantly affected by the context in which learning takes place (Chong, 2020; Han, 2019). Sociocultural theory and ecological perspective take a sociocognitive stance and perceive learning as "emergent and situated" (Han & Hyland, 2019). Furthermore, as van Lier (1997) argues, the interconnectedness of the affordances of the learning condition with the mind of the learner forms a complex adaptive system, which can affect learners' learning process and product (Kramsch, 2008). Similarly, mediation, proposed in sociocultural theory, accentuates the significant role of meditational tools in the regulation of learners' cognition and the consequent behaviors. The findings of the present study, by showing how synchronous and asynchronous peer feedback conditions resulted in different levels of peer feedback accuracy, provide further empirical evidence for these theories in an L2 learning context. Based on the findings of this study, the context in which learners provide peer feedback can provide them with different affordances, which can influence the quality of their products (accuracy of their comments).

Providing peer feedback is a cognitively difficult task as students are required to employ their knowledge and skills to review, clarify, and correct others' texts (Ballantyne, Hughes, & Mylonas,

2002; Hovardas et al., 2014). One of the reasons that might have contributed to the significant difference in the number of accurate comments in the asynchronous and synchronous conditions can be the amount of cognitive load these two different computer-mediated conditions impose. While the students in the asynchronous group had time to iteratively read and analyze the texts, formulate their comments, revise them, share them with their peers and finally engage in discussions, the students in the synchronous group had to read the texts, analyze them, provide comments, and discuss them under time constraints, which could increase the cognitive demand of the task and lead to EFL students' failure to identify errors and/or result in an increase in the number of inaccurate comments.

Cognitive load theory assumes that working memory, which is limited in nature, can be under pressure by the complexity of the information (intrinsic load), individual learning characteristics (germane load), and instructional procedure (extraneous load) (Shepherd & Bolliger, 2011). In the peer feedback accuracy case, the asynchronous computer-mediated condition seems to be superior for two reasons. The first advantage resides in providing a better processing condition (i.e., absence of temporal pressure) to reduce the extraneous load of the task. As Lee and Kalyuga (2014) argue, cognitive resources consumed by extraneous load might leave insufficient resources for dealing with intrinsic load. This can lead to lower performance (which is fewer accurate comments in a synchronous group). Furthermore, prior studies (Opfermann, Scheiter, Gerjets, & Schmeck, 2013; Woolfolk, 2016) have shown that low levels of extraneous load can optimize the germane cognitive load. The asynchronous condition (imposing lower extraneous load because of lower time pressure) could provide a better condition for all L2 learners, especially those with low working memory, to review the texts, give comments, and possibly clarify them under less cognitive load. On the other hand, the condition provided in synchronous computer-mediated feedback, which requires learners to perform all input and output processing in the working memory under immense temporal pressure, can tax EFL learners' abilities. In second language writing literature, Lee (2017) and Ferris (2010) have argued for the superiority of written feedback over oral one which is similar to synchronous feedback (e.g., Lamy & Hampel, 2007) because of the former's lower cognitive demand, especially on less-proficient learners.

Similarly, one of the findings of this study was the significantly higher number of accurate comments by low-intermediate students in the asynchronous condition. This is in line with the findings of prior studies (Lin & Griffith, 2014; Shang, 2017), which found that low-intermediate students' limited L2 knowledge might adversely affect their performance in a synchronous peer feedback condition. The lower cognitive load imposed by the asynchronous condition can help low-intermediate L2 students, who are of lower language ability, identify errors and formulate their comments more accurately as they perform different stages sequentially, with the slightest temporal pressure. This finding is in line with that of Pollock, Chandler, and Sweller (2002), which demonstrated that novice learners benefit from a design in which the researcher initially presented complex materials as a sequence of smaller isolated units of information. This sequencing is similar to the process of peer feedback provision in the asynchronous computer-mediated condition.

The second research question of this study addressed the relationship between peer feedback types (referentially-justified comments versus non-referentially-justified comments) and peer feedback accuracy. The analysis of the data indicated that only one-third of the comments were referentially-justified, but they were significantly more accurate than non-referentially-justified comments. As the findings of a prior study (Hoomanfard & Rahimi, 2020) showed, EFL students employ references to external resources to benefit from the authority of acceptable resources. Furthermore, they assign more credibility to those comments which are generated by valid sources such as books, teachers, or parents (Brown et al., 2009; Strijbos et al., 2010). Possibly, considering the positive perceptual effects of dependable sources on their peers, the EFL students employed references to valid resources to justify their own comments. The by-product of this strategy, as the findings showed, was the significantly higher level of peer feedback accuracy.

However, a possible reason for the difference in accuracy levels of comments provided by the students in these two contexts might reside in the activities involved in the formulation of these two feedback types. In the case of referentially-justified peer feedback, students' search for a suitable reference to support their own suggestion or provision of a suggestion as the result of finding the proper reference can help the students involve in constructive investigations (Thomas, 2000), which due to their depth of engagement with L2 sources can decrease the chances of inaccurate comments. Although project-based learning activities include the co-construction of knowledge among peers in the knowledge formation stage, in referentially-justified peer feedback instances, the co-construction of knowledge is mediated by an artifact such as a book, a text, or a teacher's speech. However, the engagement of L2 learners with these sources seems to be curbed by the time constraints imposed by a synchronous computer-mediated condition, which, in turn, might lead to more inaccuracies in the provided comments.

Conclusions, Implications, and Future Research

The present study contributes to a growing but still relatively small body of research on peer feedback accuracy. The findings regarding the first research question of this study acknowledge the significance of the feedback exchange condition in improving the accuracy of peer comments provided by L2 intermediate students. To be more specific, the results indicate the superiority of an asynchronous computer-mediated condition as it imposes a lower cognitive load on learners. This can help L2 intermediate students pinpoint more errors and provide more accurate comments. This finding provides further evidence to support the significant role of contextual factors in the success of written feedback exchange activities (e.g., Alshuraidah & Storch, 2019; Chang, 2009, 2012; Ellis, 2010; Hoomanfard, Jafarigohar, Jalilifar, & Hosseini, 2018; Liu & Sadler, 2003; Rahimi & Zhang, 2016; Wang, 2009). Although in his study, Rahimi (2013) indicated that peer feedback training (leading to students' higher levels of germane load) could improve the accuracy of peer comments, the findings of the present study showed that even in those cases that EFL students are well-trained, the condition in which peer feedback comments are exchanged can influence students' comments and their accuracy.

While some researchers have recommended synchronous peer feedback for being more interactive (Braine, 2001) and providing immediate input (Shang, 2017), based on the findings of the present study, the researchers would like to suggest that L2 language teachers should implement asynchronous peer feedback when the accuracy of comments is their priority; however, the combination of asynchronous and synchronous conditions can help students, especially the less linguistically competent ones, to provide quality feedback and benefit from the interactive discussions. While this study showed the superiority of asynchronous computer-mediated conditions in producing more accurate comments, further studies can be conducted to uncover the extent to which temporal constraints (synchronous condition) can affect L2 learners' error detection performance.

Further, based on the findings, L2 students provide significantly more referentially-justified comments in the asynchronous computer-mediated condition. These comments were significantly more accurate than those not supported by references to credible external sources such as textbooks, online sources, and teachers' lectures. Prior studies conducted by Narciss (2008), Strijbos et al. (2010), and Walker (2015) have shown that the content of peer feedback can affect the students' attitudes, feedback acceptability, and revisions performance. The significantly higher accuracy level of referentially-justified peer comments provides evidence that the content of peer feedback can also have a relationship with its accuracy. Although further research is needed on referentially-justified peer feedback in various contexts, results of this study suggest that second

language teachers should ask their students to use referentially-justified peer feedback in L2 writing courses to improve the accuracy level of comments provided by them.

The researchers believe that involving L2 students in providing referentially-justified peer feedback can benefit feedback givers as they will engage more deeply with the process of providing comments and can strengthen their writing ability (including language use, vocabulary, content, organization, and mechanics) in the short and long-run. In addition, it can improve feedback receivers' attitudinal, behavioral, and cognitive engagement with peer feedback. However, empirical research is needed to prove these claims; thus, further studies can investigate whether providing referentially-justified peer feedback can benefit feedback providers in expanding their L2 writing ability significantly more than non-referentially-justified peer comments. On the feedback receiver side, further studies can be conducted to examine the effects of using referentially-justified peer feedback on receivers' peer feedback perceptions and peer feedback rate of incorporation. In addition, using eye-tracking technology, other researchers can compare the cognitive processes that students receiving referentially-justified and non-referentially-justified comments go through to apply their peers' comments. Finally, the effect of implementing this feedback type on L2 learners' writing ability in academic writing contexts can be investigated.

References

Alqassab, M., Strijbos, J. W., & Ufer, S. (2019). Preservice mathematics teachers' beliefs about peer feedback, perceptions of their peer feedback message, and emotions as predictors of peer feedback accuracy and comprehension of the learning task. *Assessment & Evaluation in Higher Education*, *44*(1), 139-154.

Alshuraidah, A., & Storch, N. (2019). Investigating a collaborative approach to peer feedback. *ELT Journal*, *73*(2), 166-174.

Ary, D., Jacobs, L. C., Irvine, C. K. S., & Walker, D. (2018). *Introduction to research in education*. *Belmont, CA:* Cengage Learning.

Bahari, A. (2021). Computer-mediated feedback for L2 learners: Challenges versus affordances. *Journal of Computer Assisted Learning*, *37*(1), 24-38.

Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education*, *27*(5), 427-441.

Berndt, M., Strijbos, J. W., & Fischer, F. (2018). Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing. *European Journal of Psychology of Education*, *33*(1), 31-49.

Bolzer, M., Strijbos, J. W., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eye-tracking: Role of feedback-characteristics, fixation-durations and transitions. *Journal of Computer Assisted Learning*, *31*(5), 422-434.

Braine, G. (2001). A study of English as foreign language (EFL) writers on a local area network (LAN) and in traditional classes. *Computers and Composition, 18*, 275-292.

Brown, G. T., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. (2009). Use of interactive–informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction*, *19*(2), 97-111.

Cha, Y. (2007). A study of peer feedback patterns in CMC modes on Korean EFL students. *Multimedia Assisted Language Learning*, *10*(1), 9-35.

Chang, F. (2009). Peer review through synchronous and asynchronous modes: A case study in a Taiwanese college English writing course. *JALTCALL Journal*, *5*(1), 45-64.

Chang, F. (2012). Peer review via three modes in an EFL writing course. *Computers and Composition, 29*, 63-78.

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, *48*(3), 409-426.

Chong, I. (2017). How students' ability levels influence the relevance and accuracy of their feedback to peers: A case study. *Assessing Writing*, *31*, 13-23.

Chong, S. W. (2020). Reconsidering student feedback literacy from an ecological perspective. *Assessment & Evaluation in Higher Education*, *46*(1), 92-104.

Ellis, R. (2010). A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, *32*(2), 335-349.

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, *32*(2), 181-201.

Gao, Y., Schunn, C. D. D., & Yu, Q. (2019). The alignment of written peer feedback with draft problems and its impact on revision in peer assessment. *Assessment & Evaluation in Higher Education*, *44*(2), 294-308.

Gedera, D. S. (2016). The application of activity theory in identifying contradictions in a university blended learning course. In D. S. P. Gedera & P. J. Williams (Eds.), *Activity theory in education* (pp. 53-69). Sense Publishers.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, *20*(4), 304-315.

Han, Y. (2019). Written corrective feedback from an ecological perspective: The interaction between the context and individual learners. *System*, *80*, 288-303.

Han, Y., & Hyland, F. (2019). Academic emotions in written corrective feedback situations. *Journal of English for Academic Purposes*, *38*, 1-13.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112.

Holmes, J. (2001). *An introduction to sociolinguistics*. Harlow: Pearson Education Limited.

Hoomanfard, M. H., Jafarigohar, M., Jalilifar, A., & Hosseini, S. M. (2018). Comparative study of graduate students' self-perceived needs for written feedback and supervisors' perceptions. *Journal of Research in Applied Linguistics*, *9*(2), 24-46.

Hoomanfard, M. H., & Rahimi, M. (2020). A comparative study of the efficacy of teacher and peer online written corrective feedback on EFL learners' writing ability. *Zaban Pazhuhi (Journal of Language Research)*, *11*(33), 327-352.

Hovardas, T., Tsivitanidou, O. E., & Zacharia, Z. C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. *Computers & Education*, *71*, 133-152.

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.

Jafarigohar, M., Hoomanfard, M. H., & Jalilifar, A. (2018). A typology of supervisor written feedback on L2 students' theses/dissertations. *Iranian Journal of Applied Linguistics*, *21*(2), 43-87.

Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: their origin and impact on revision work. *Instructional Science*, *39*(3), 387-406.

Kim, S. H. (2015). Preparing English learners for effective peer review in the writers' workshop. *The Reading Teacher*, *68*(8), 599-603.

Kramsch, C. (2008). Ecological perspectives on foreign language education. *Language Teaching*, *41*(3), 389.

Kumar, V., & Stracke, E. (2007). An analysis of written feedback on a PhD thesis. *Teaching in Higher Education*, *12*(4), 461-470.

Lamy, M. N., & Hampel, R. (2007). *Online communication in language teaching and learning*. Australia: Palgrave Macmillan.

Lantolf, J. P. (2011). The sociocultural approach to second language acquisition: Sociocultural theory, second language acquisition, and artificial L2 development. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 36-59). Routledge.

Lee, C. H., Kalyuga, S. (2014). Expertise reversal effect and its instructional implications. In V. A. Benassi, C. E. Overson, C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into curriculum* (pp. 31-44). American Psychological Association - STP.

Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer.

Leki, I. (1990). Potential problems with peer responding in ESL writing classroom. *CATESOL Journal, 3*, 5-17.

Lin, J.Y. (2005). Synchronous and asynchronous conferencing: A comparison of two modes of online ESL peer response and their effects on student talk and subsequent text revision. *Dissertation Abstracts International, 66*(7), 25-48.

Lin, S., & Griffith, P. (2014). Impacts of online technology use in second language writing: A review of the literature. *Reading Improvement, 51*(3), 303-312.

Liu, J., & Hansen, J. G. (2018). *Peer response in second language writing classrooms*. University of Michigan Press: Ann Arbor.

Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, *2*, 193-227.

Liu, N. F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher education*, *11*(3), 279-290.

Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. Van Merrie¨nboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed.). (pp. 125-143). Erlbaum.

Nguyen, T. (2018). The effect of combined peer-teacher feedback on Thai students' writing accuracy. *Iranian Journal of Language Teaching Research*, 6(2), 117-132.

Opfermann, M., Scheiter, K., Gerjets, P., & Schmeck, A. (2013). Hypermedia and self-regulation: An interplay in both directions. In *International handbook of metacognition and learning technologies* (pp. 129-141). Springer.

Panadero, E., & Brown, G. T. (2017). Teachers' reasons for using peer assessment: positive experience predicts use. *European Journal of Psychology of Education*, *32*(1), 133-156.

Panadero, E., Romero, M., & Strijbos, J. W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, *39*(4), 195-203.

Patchan, M. M., Hawk, B., Stevens, C. A., & Schunn, C. D. (2013). The effects of skill diversity on commenting and revisions. *Instructional Science, 41,* 381–405.

Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feedback: How students respond to peers' texts of varying quality. *Instructional Science*, *43*(5), 591-614.

Pham, H. T. P. (2020). Computer-mediated and face-to-face peer feedback: Student feedback and revision in EFL writing. *Computer Assisted Language Learning*, 1-37.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, *12*(1), 61-86.

Rahimi, M. (2013). Is training student reviewers worth its while? A study of how training influences the quality of students' feedback and writing. *Language Teaching Research*, *17*(1), 67-89.

Rahimi, M. (2015). The role of individual differences in L2 learners' retention of written corrective feedback. *Journal of Response to Writing*, *1*(1), 19-48.

Rahimi, M., & Zhang, L. J. (2016). The role of incidental unfocused prompts and recasts in improving English as a foreign language learners' accuracy. *The Language Learning Journal*, *44*(2), 257-268.

Rastgou, A., Storch, N., & Knoch, U. (2020). The effect of sustained teacher feedback on CAF, content and organization in EFL writing. *Iranian Journal of Language Teaching Research*, 8(2), 41-61.

Rotsaert, T., Panadero, E., Estrada, E., & Schellens, T. (2017). How do students perceive the educational value of peer assessment in relation to its social nature? A survey study in Flanders. *Studies in Educational Evaluation*, *53*, 29-40.

Sánchez-Naranjo, J. (2019). Peer review and training: Pathways to quality and value in second language writing. *Foreign Language Annals*, *52*(3), 612-643.

Shang, H. F. (2017). An exploration of asynchronous and synchronous feedback modes in EFL writing. *Journal of Computing in Higher Education*, *29*(3), 496-513.

Shepherd, C. E., & Bolliger, D. U. (2011). The effects of electronic portfolio tools on online students' perceived support and cognitive load. *The Internet and Higher Education*, *14*(3), 142-149.

Shintani, N. (2015). The effectiveness of processing instruction and production-based instruction on L2 grammar acquisition: A meta-analysis. *Applied Linguistics*, *36*(3), 306-325.

Sluijsmans, D. M., Brand-Gruwel, S., & van Merriënboer, J. J. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, *27*(5), 443-454.

Stracke, E., & Kumar, V. (2010). Feedback and self-regulated learning: Insights from supervisors' and PhD examiners' reports. *Reflective Practice*, *11*(1), 19-32.

Strijbos, J. W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency? *Learning and Instruction*, *20*(4), 291-303.

Thomas, J. W. (2000). *A review of research on project-based learning.* Autodesk Foundation.

Tian, L., & Li, L. (2018). Chinese EFL learners' perception of peer oral and written feedback as providers, receivers and observers. *Language Awareness*, *27*(4), 312-330.

Van der Kleij, F. M., & Lipnevich, A. A. (2020). Student perceptions of assessment feedback: A critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability, 33*(2), 1-29.

Van Lier, L. (2004). The semiotics and ecology of language learning. *Utbildning & Demokrati*, *13*(3), 79-103.

Van Zundert, M. J., Könings, K. D., Sluijsmans, D. M. A., & Van Merriënboer, J. J. G. (2012). Teaching domain-specific skills before peer assessment skills is superior to teaching them simultaneously. *Educational Studies*, *38*(5), 541-557.

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.

Ware, P. & O'Dowd, R. (2008). Peer feedback on language form in telecollaboration. Language Learning & Technology, 12(1), 43-63.

Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal, 13*(2), 7–26.

Woolfolk, A. (2016). *Educational Psychology*. Pearson Education, Inc.

Yeh, S. W., Lo, J. J., & Chu, H. M. (2014). Application of online annotations to develop a web-based Error Correction Practice System for English writing instruction. *System*, *47*, 39-52.

Zhao, H. (2010). Investigating learners' use and understanding of peer and teacher feedback on writing: A comparative study in a Chinese English writing classroom. *Assessing Writing, 15*, 3–17.

Appendix A

Writing Tasks

You have 40 minutes to write at least 250 words to argue the following topic.

Do you agree or disagree with the following statements? People should sometimes do things they do not enjoy doing. Use specific reasons and examples to support your answers.

You have 40 minutes to write at least 250 words to argue the following topic.

Some people prefer to live in a small town. Others prefer to live in a big city. Which place would you prefer to live in? Use specific reasons and details to support your answer.

Appendix B

The Process of Categorizing Comments based on Speech Functions

The researchers initially categorized the comments into the three speech function categories provided by Holmes (2001). This categorization posits that human interactions can be put in one of the categories of directive (utterances attempting to get someone to do something), expressive (utterances expressing the speaker's feelings), and referential (utterances providing information). For instance, a comment like *I like your introduction*. can be coded as expressive; *You have to write more examples to support your topic sentence*. can be coded as directive, and *The topic sentence at the end of the paragraph? I know a topic sentence can be put at the beginning, at the end, or in the middle of a paragraph. But, when you start your paragraph with your topic sentence, your reader has an easier task to follow what you want to establish*. can be assigned to the referential comments. The provided comments were coded into these three categories, and in order to examine the accuracy of comments, the researchers excluded the expressive comments from the corpus since they represented the writer's feelings, which could not be evaluated as accurate or inaccurate as they showed the feelings of the feedback providers. The accuracy of directive and referential comments was also scrutinized.

In order to make sure about the coding process, several procedures were undergone. First the researchers coded comments into the three speech functions. The lead author examined all comments and coded them into referential, expressive, and directive categories. In this study referential and directive comments were included. A PhD holder of TEFL (Teaching English as a Foreign Language), who was familiar with the categories, coded half of the comments and the obtained inter-coder (Cohen's Kappa) coefficient was .96. The discrepancies, which were mostly about the indirect function of comments, were discussed to reach unanimous decisions by the three researchers and the guest scholar. To give an example, a comment like *"In English, we use -ing form of the verb after prepositions."* which was accompanied by underlining a faulty sentence was coded by one of the researchers as referential and as directive, as it indirectly requested the writer to change the form of the verb. In this case, the comment was coded as both referential and directive.

Saeed Mehrpour is a professor in ELT at Istanbul Aydin University, Turkey. He has published extensively in national and international journals over the past 18 years. His main research interests encompass writing, reading, and language teacher education.

Mohammad Hamed Hoomanfard has received his PhD in TEFL from Payame Noor University, Tehran, Iran. He has published more than 20 papers in national and international journals. His main research interests are writing, feedback, and critical pedagogy.

Elham Vazin holds an MA in TEFL from Shiraz University. She is interested in conducting research on writing.