Journal of Language Horizons

Scientific Quarterly Journal of Language Horizons, *Alzahra University*
Volume 8, Issue 1, Spring 2024
Research Article
pp. 139-163

TEFL Master's Degree Students' Incorporation of Peer Feedback in Academic Texts

Saeid Bahadoranfar[1]
Manoochehr Jafarigohar[*2]
Hassan Soleimani[3]

Received: 31/10/2022 Accepted: 11/12/2023

Abstract

The examination of the literature reveals that scant attention has been paid to L2 graduate students' engagement with feedback on their academic texts. To fill this gap in the literature, the current study investigated the engagement of 53 TEFL master's degree students with peer comments using a host of data collection and analysis tools, including eye-tracking, stimulated recall interviews, and content analysis. The participants exchanged computer-mediated comments on an academic writing task (i.e., thesis proposal) and were asked to revise their texts based on these comments while their eye movements were being captured using an eye-tracking application. Then, the eye-tracking videos were employed as stimuli to extract the participants' reasons for not applying the comments. In addition, the participants' first and revised texts underwent content analysis, and their feedback incorporation strategies were identified. The findings of this study indicated that the participants applied more than two-thirds of the comments; however, justified elaborated feedback was adopted more than elaborated and concise general feedback. The participants also applied justified elaborated feedback more accurately than the other feedback types. Finally, the results showed that four main feedback qualities (being too general, incomprehensible, inapplicable, and faulty) adversely affected L2 MA students' feedback incorporation.

Keywords: academic writing, engagement, feedback incorporation, peer feedback, second language writing

* Corresponding author

[1] Department of TEFL and English Literature, Payame Noor University, Tehran, Iran; saeidbaha84@student.pnu.ac.ir

[2] Department of TEFL and English Literature, Payame Noor University, Tehran, Iran; jafari@pnu.ac.ir

[3] Department of TEFL and English Literature, Payame Noor University, Tehran,Iran; h_soleimanis@pnu.ac.ir

DOI: 10.22051/lghor.2023.41413.1751

Introduction

One of the factors that can affect the success of feedback on non-native students' academic texts is the extent to which learners engage with the provided feedback to both learn a new item and modify their texts while they are fully motivated to accomplish their task (Berndt et al., 2018). Thus, unlike the mainstream trend in second language writing studies, which chiefly focuses on the *providing* side of the feedback exchange, the study of learners' engagement with feedback mainly emphasizes the role of *receiving* side of the feedback exchange in the success of feedback activities.

While several studies have investigated L2 (referring to both foreign and second language) learners' engagement with feedback in general English and undergraduate contexts, L2 graduate students' engagement with their peer feedback has still remained an underexplored research area, and empirical studies are required to uncover to understand how graduate students engage with feedback on their academic texts. Furthermore, eye-tracking technology that can provide precise data on writers' engagement with feedback has been used minimally in second/foreign language learning contexts.

As an attempt to bridge a part of these gaps in the literature, the present study focuses on the examination of Iranian L2 master students' engagement with peer feedback on their thesis proposals by analyzing their incorporation strategies, the accuracy of their revision, and reasons for not applying their peer comments. To collect accurate data, the researchers recorded the participants' incorporation process using an eye-tracking application and used them in stimulated recall interviews to uncover the learners' reasons for not engaging with the incoming comments.

Literature Review

The importance of learner engagement with feedback has been well-documented in the literature on second language writing. Behavioral engagement is reported to determine the success of a feedback practice and explain the differential success of students receiving comments in second language writing programs (Han, 2017; Han & Hyland, 2019). The examination of empirical studies on students' engagement with feedback reveals that students' poor engagement with feedback results in their failure to learn from comments (Sinclair & Cleland, 2007).

In recent years, some scholars (Han & Hyland, 2019; Yuan & Kim, 2018; Zhang & Hyland, 2018) have attempted to redefine the concept of engagement to match it with feedback on students' written products. Zhang and Hyland (2018) asserted that the three components of learner engagement were relevant to feedback activity. For instance, emotional (affective) engagement deals with learners' attitudinal reactions to feedback. The extent to which learners feel frustrated, stressed, or motivated determines learners' emotional engagement with feedback. The literature on feedback on L1 and L2 writing literature has accommodated studies (e.g., Donia et al., 2022; Mulliner & Tucker, 2017; Ryan & Henderson, 2018; Zhan et al., 2022) examining learners' perceptions and attitudes toward teacher and peer feedback. Cognitive engagement refers to "how students attend to feedback encompassing the use of revision operations (strategies) and cognitive (metacognitive) strategies" to uptake the provided items (Zhang & Hyland, 2018, p. 9). Yuan and Kim (2018) argued that within feedback activities, students' behavioral engagement equals the extent to which they examine their peers' or teachers' comments and incorporate them into their texts.

Studies on Learners' Engagement with Feedback

Some studies have investigated the reasons that prevented graduate students' engagement with feedback. For instance, the results of the mixed-methods study conducted by Carless (2006) showed that one of the significant reasons that hindered students' incorporation of comments was their inability to comprehend them. Carless (2006) maintained that students' lack of academic discourse knowledge could lead to this incomprehensibility, which can affect the revision pattern. Similarly, Sadler (2010) admitted the significant role of learners' self-perceived ability to apply comments in forming students' level of engagement. Hoomanfard and Rahimi (2020) found that learners with higher English language proficiency levels were more successful in understanding comments and incorporating them into their revised versions.

Another set of studies has examined how feedback type can affect L2 writers' engagement with comments. 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. The research carried out by Bai and Hu (2016) also showed the effect of feedback type on second language learners' incorporation patterns. Another study by Berndt et al. (2018) investigated the effect of peer feedback content (concise general vs. elaborated specific feedback) and sender's competence on students' perceptions, revision performance, and mindful cognitive processing. In a more recent study, Mohammed and Al-Jaberi (2021) investigated graduate students' engagement with feedback and found that feedback type could determine their feedback incorporation.

Eye-Tracking Technologies in Feedback Studies

Few researchers have used eye-tracking technologies to examine learners' responses to written feedback on their texts. Eye-tracking data have been used to understand whether different feedback types made a difference in learners' acquisition of grammatical items. For instance, Shintani and Ellis (2013) benefited from eye-tracking data to examine the extent to which direct written corrective feedback and metalinguistic explanation affected second language learners' explicit and implicit knowledge of a specific grammatical item in English. Eye-tracking has also been employed to uncover the effects of different feedback types on learners' behavioral engagement with feedback. Bolzer et al. (2015) and Berndt et al. (2018) examined the performance of tertiary level students who were assigned to elaborated specific and elaborated specific plus justification feedback groups and identified how the change in the feedback influenced learners' behavioral engagement with feedback. Ranalli (2021) also found that specific comments were significantly more effective than general comments in engaging learners.

The Present Study

Although the literature on second language learning has emphasized the significance of learner engagement with feedback (Han & Hyland, 2019; Ranalli, 2021), this area has remained an underexplored one (Amiryousefi, 2019; Aubrey et al., 2020; Stevenson & Phakiti, 2019). The examination of the literature shows that most previous studies have focused on the product of learners' engagement with

supervisor (instructor) feedback, and few recent studies (Mohammed & Al-Jaberi, 2021; Ranalli, 2021) have examined L2 graduate students' behavioral and cognitive engagement with peer feedback on their academic texts.

Furthermore, the number of studies using eye-tracking technologies to examine feedback is limited to a few (Berndt et al., 2018; Bolzer et al., 2015; Cutumisu et al., 2019; Ranalli, 2021; Shintani & Ellis, 2013). Eye-tracking technology is significant since the data obtained from this data collection option can provide researchers with accurate information about feedback receivers' cognitive processes (Ranalli, 2021).

The present study aimed to fill a part of the mentioned gaps in the literature by using content analysis to examine how L2 master's degree students use peer comments on their thesis proposals. The participants' revision accuracy was also examined in this research. This study also investigates the extent to which L2 MA students apply their peers' comments, and their reasons for not applying the comments are studied using eye-tracking technology and stimulated recall interviews. The following questions guided this study.

1. To what extent do L2 master's degree students incorporate peer comments into their thesis proposals?
2. What reasons prevent L2 master's degree students from applying peer comments?
3. Does different peer feedback type affect L2 master's degree students' accuracy of feedback incorporation?

Method

Participants and Corpus

The sample consisted of 53 TEFL master's degree students who were required to write research proposals in their Academic Writing course. The participants attended two classes (*N1* =23 and *N2* =30) at a university in Shiraz, Iran. They were selected based on a convenience sampling procedure. Both male (*N*= 24) and female (*N* = 29) students participated in this study, and their ages ranged between 23 and 34 (*M* = 26.4, *SD* = 3.1). Only two students had published papers in national peer-reviewed journals before the data collection started, and the rest of the participants did not have any academic texts published. All participants had already

worked with Microsoft Word to write academic texts and exchange comments. However, a link to a 30-minute video showing how to use Microsoft Word was sent to all users so that they learn how to use Word to exchange comments.

The corpus of this study included 53 research proposals (maximum 1200 words excluding references). The mean length of texts was 1413 words (*SD* = 98 words). These proposals were prepared by the participants to be submitted to their thesis supervisors. All proposals included introduction, literature review, and method sections. The revised (second) drafts of these proposals and the peer comments were other parts of the studied corpus. The details about the provided comments are provided in the Results Section.

Instruments

Eye-Tracking Program

To capture the participants' eye movements, the researchers used *GazeRecorder* software. Unlike other options, which required the participants to come to a data collection site, GazeRecorder works with the webcam of the participants' laptop or desktop computers. This feature improved the quality of data collection since it was done less obtrusively since no extra device was attached to their computers, and they applied the comments at their preferred place. This program tracks the eye movements of each person and can show the areas which were focused on for a longer time. Figure 1 shows an image of the eye-tracking data. A short video of the eye-tracking data collection is also accessible via this link to short a video. The colors recorded showed the amount of time each learner spent on a specific area of the text. While blue and green showed a low amount of time, yellow and red points were indicatives of a higher amount of fixation on an area.

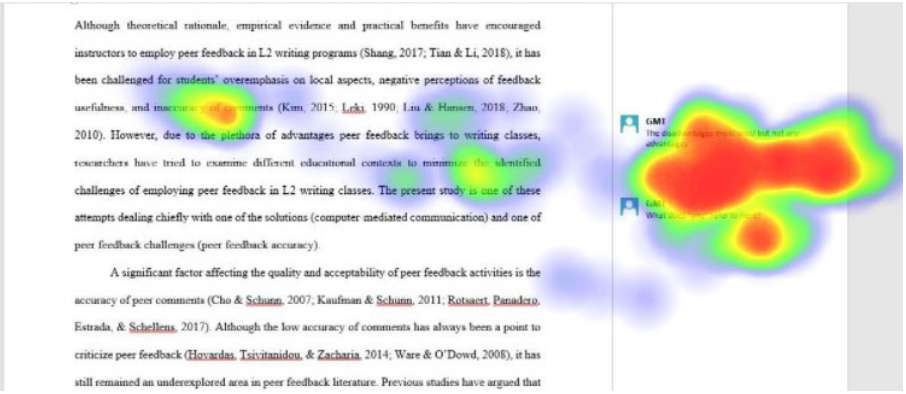
Stimulated Recall Interview

The researchers used stimulated recall interviews to examine the participants' reasons for not incorporating peer comments. The researchers used the videos from the eye-tracking procedure to stimulate learners to talk about their thoughts while applying their peer comments. The videos were played right after the participants finished applying the comments. The participants were asked to talk about their thoughts as they watched the videos of their eye movements. The

researchers asked about the reasons for fixations, pauses, and revisions. These interviews were audio-recorded and transcribed for further analysis. In stimulated recall interviews, which intend to uncover cognitive processes, simple questions such as *Why did you do so? What were you thinking about then? What were you doing then?* are asked (Godfroid & Hui, 2020). The same questions were employed in this study to uncover the participants' thoughts while applying the comments.

Figure 1

A Sample of Eye-tracking Data



Procedure and Data Collection

The data required to answer the research questions were collected during the Covid-19 Pandemic; thus, all activities involved in this study were incorporated into the present computer-mediated classes. This condition facilitated the data collection process for the researchers, who could participate in all sessions as non-participant observers. The data of this study were collected in two online academic writing classes at a university in Shiraz, Iran. In the first six sessions, the participants were given lectures on how to write a thesis proposal. Meanwhile, the participants were provided with high-quality peer feedback provision and incorporation features. They were also required to accomplish minor writing tasks, provide peer feedback on their peers' texts, and finalize their thesis topics with their instructors.

As the main assignment of their course, the participants had to write proposals (maximum 1200 words excluding references) and submit them by the 14th session. These proposals were reviewed by the participants' peers from the same

classes. The participants were paired by the instructor so that they could exchange their texts and comments. The participants had to select the topics of their proposals. The participants were required to provide comments in a week and send the comments to the instructor. Figure 2 provides a schematic summary of the data collection procedure. Although instructor feedback was also used in the course, its data were not included in our study since this feedback type was out of the scope of the present research.

Figure 2

Data Collection Steps

Based on an agreed-upon schedule, the participants went online for the feedback incorporation task. In this data collection phase, one of the researchers (who was not the instructor of the classes) handled the feedback incorporation sessions. First, the participants calibrated the eye-tracking system using their webcams. Then, they were shown the commented-on drafts. The comments were in the form of Track Changes, marginal comments, or general concluding comments. The students had 150 minutes to read the comments, decide whether and how to apply them, and revise their texts. The participants' eye-tracking data were recorded in cloud storage. On average, each student spent 78 minutes (SD = 8.63) revising their papers. Immediately after each feedback incorporation session, the stimulated recall interviews were initiated. Each student was shown the video, and she/he was asked about her/his thoughts or emotions. The researcher interrupted the silence for more than 30 seconds with a question asking about the interviewee's thoughts. The researchers informed the participants that the screen content and their voice were recorded before starting the feedback incorporation activity to observe the ethical considerations.

Data Analysis

The comments provided by Iranian L2 master's degree students were deductively categorized into three main categories (Table 1), which were taken from the related literature in general education (Berndt et al., 2018; Bolzer et al., 2015; Walker, 2015).

Table 1

Feedback Types Based on Feedback Specificity

Feedback type	Concise general feedback	Elaborated feedback	Justified elaborated feedback
Description	Mention the problem in the text with no further information	Provide not only the position and error type but information on how to proceed to solve the problem	Provide an explanation of a correction to a content or skills shortcoming, or an explanation of why something is praiseworthy.
Feedback examples	Your conditional sentences are not correct.	Use were instead of are in the If clause	It is conditional type II, and you should use past tense in the if-clause.
	You have to elaborate on your ideas more.	Explain your idea and provide examples.	To have a persuasive paragraph, you need to support your topic sentence using explanations and examples.

As Table 1 shows, the three feedback types included comments ranging from the most general ones (concise general feedback) to justified elaborated feedback, which is the most detailed one. The comments provided by the study participants were first categorized by one of the researchers in this study. Then, an assistant professor in the field of applied linguistics categorized half of these comments deductively, and the inter-coder reliability of .96 was obtained. The few disagreements were discussed carefully until unanimous decisions were made.

To answer the first research question, the researchers analyzed the first and

revised versions as well as comments on proposals to determine the percentages of applied comments, ignored comments, and text modification to avoid feedback incorporation. This part of the analysis was carried out in a number of steps. First, the researcher excluded those comments that only included feedback providers' positive opinions about the text since these comments did not require modification. All directive (including requests, questions, and orders) and negative expressive comments (implying changes in the text) were studied to check if they were applied. In those cases where both expressive and directive comments addressed a single issue, only one of them was counted.

To ensure the credibility of the coding process, several measures were taken. First, a researcher and a TEFL associate professor out of the current research team examined 50 percent of the comments to code them deductively as applied, ignored, and text modification to avoid feedback incorporation, and the inter-coder reliability of .92 was obtained. All of the discrepancies were related to the comments on global aspects of writing (content and organization). These comments were analyzed in an extensive meeting until unanimous decisions were made.

The second question was answered by analyzing the participants' reasons for not applying the comments: a) the stimulated recall interview data were transcribed carefully; b) one of the researchers categorized the statements that showed their reasons for not applying the comments inductively. Four main categories were identified. The same coder, introduced above categorized one-third of the statements deductively and the inter-coder reliability (Cohen's Kappa) was .96. The discrepancies were discussed until both parties reached a full agreement. Next, c) a member check was done by asking four participants to code one-third of the statements into the categories deductively, and the inter-coder reliability values of .97, .98, .97, and .95 were achieved. To reach referential adequacy, the researchers provided some direct quotations for each category to elucidate the issues for the audience. It should be noted that the researchers of this study were not the instructors of the examined classes, and the participants were informed that their cooperation in this research could not affect their scores.

To answer the third research question, which addressed the effect of feedback type on L2 students' accuracy of feedback incorporation, the researchers examined the first versions, second versions, and comments to examine the extent to

which comments had been incorporated accurately. The analysis of form-related comments was straightforward. One of the researchers and an external reviewer, an experienced English language instructor with a PhD in TEFL, examined the accuracy of the revisions required by form-related comments, and the inter-coder reliability of .93 was achieved. However, since examining the revision accuracy of content-related comments was subjective, two researchers of this study examined the revisions individually and reached the inter-coder reliability of .86. After discussing the discrepancies and reaching unanimous decisions, the instructor (mentioned in this paragraph) examined half of the comments, and the inter-coder reliability was .96. The three researchers attended an online meeting and reached unanimous decisions for the remaining items.

Results

The examination of the data showed that a total of 1402 comments were left on the participants' texts. On average, they received 32.11 (*SD* = 6.86) comments on their texts each. The mean length of texts was 943 (*SD* = 49.3) words, and the mean length of comments was 14.2 (*SD* = 3.6) words. The comments provided by the students addressed both form-related (798, 56.92 %) and content-related (*N*= 604, 43.08 %) issues. The analysis of the provided comments indicated how the provided comments were distributed in the three feedback types and five major feedback focuses (Table 2).

Table 2

Feedback Types Provided by Master's Degree Students

	Concise General Feedback	Elaborated Feedback	Justified Elaborated Feedback	Total
Grammar	110 (32.64%)	146 (43.32%)	81 (24.03%)	337
Mechanics	98 (36.29%)	99 (36.66%)	73 (27.03%)	270
Organization	106 (33.65%)	124 (39.36%)	85 (26.98%)	315
Content	115 (39.79%)	120 (41.52%)	54 (18.68%)	289
Vocabulary	59 (30.89%)	97 (50.78%)	35 (18.32%)	191
Total	488 (34.8%)	586 (41.79%)	328 (23.39%)	1402

As Table 2 shows, the most frequent feedback type was elaborated feedback. The participants provided 586 elaborated comments on their peers' texts, accounting for 41.79 percent of all comments. The second frequent feedback type was concise general feedback (*N* = 488, 34.80 %), and the least frequent feedback type was justified elaborated feedback (*N* = 328, 23.39 %). This table also presents how different writing aspects were commented on using these three main feedback types.

Table 3

Incorporation Strategies Based on Feedback Types

	Applied	Text modified	Ignored	Total
Concise general feedback	276 (56.54%)	28 (5.73%)	184 (37.7%)	488
Elaborated feedback	373 (63.66%)	76 (12.96%)	137 (23.38%)	586
Justified elaborated feedback	287 (87.5 %)	15 (4.57%)	26 (7.92%)	328
	936 (66.76%)	119 (8.49%)	347 (24.75%)	1402

Feedback Incorporation

To examine the participants' behavioral engagement with comments, the researchers categorized the comments under three categories: applied, text modified without applying the comment, and ignored. First, the participants' behavioral engagement was examined based on feedback types (Table 3 above).

As indicated in Tables 3 and 4, the participants applied two-thirds of the comments (*N* = 936, 66.76%), and 65.27 (*N* = 611) of these incorporated comments were applied accurately. A more detailed analysis of the data showed that the highest level of feedback incorporation accuracy belonged to justified elaborated comments (77.7%), followed by elaborated feedback (61.93%) and concise general feedback (56.88%).

Table 4

Accuracy of Feedback Incorporation

	Incorporated					
	Correct			Incorrect		
	*CGF	EF	JEF	CGF	EF	JEF
Grammar	32 (29.09)	67 (45.89)	57 (70.37)	26 (23.63)	46 (31.5)	19 (23.45)
Mechanics	34 (34.69)	36 (36.36)	51 (69.86)	28 (28.57)	24 (24.24)	12 (16.43)
Organization	36 (33.96)	53 (42.74)	62 (72.94)	27 (25.47)	33 (31.13)	13 (15.29)
Content	43 (37.39)	42 (35)	32 (59.25)	30 (26.08)	28 (23.33)	12 (22.22)
Vocabulary	12 (20.33)	33 (34.02)	21 (60)	8 (13.55)	11 (11.34)	8 (22.85)
Total	157 (56.88%)	231 (61.93%)	223 (77.7%)	119 (43.11%)	142 (38.07%)	64 (22.29%)

*CGF= concise general feedback, EF= elaborated feedback, JEF= justified elaborated feedback

The data provided in Tables 3 and 5 show that the participants ignored 347 peer comments (24.75%) and modified their texts to avoid feedback incorporation (*N* = 119, 8.49 %). The results also showed that the feedback type affected the incorporation strategies. The data analysis also indicated that the participants ignored more than two-thirds of concise general comments (*N* = 184, 37.7 %), followed by 23.38 percent of unincorporated elaborated feedback. The lowest rate of ignored comments belonged to the justified elaborated feedback condition, where 7.92 percent of comments (*N* = 26) were not applied by the students. The last incorporation strategy was text modification. The highest text modification level to avoid feedback incorporation was recorded for the elaborated comments (*N* = 76, 12.96 %). The lowest rate belonged to the justified elaborated feedback (*N* = 15, 4.57 %), and concise general feedback resulted in text modification in 5.73 percent of cases (*N* = 28).

Table 5

Feedback Unincorporated

	Text Modified			Ignored		
	CGF*	EF	JEF	CGF	EF	JEF
Grammar	8 (7.27%)	16 (10.95%)	2 (2.46%)	44 (40%)	18 (12.32%)	3 (3.7%)
Mechanics	7 (7.14%)	12 (12.12%)	3 (4.1%)	29 (29.59%)	27 (27.27%)	7 (9.58%)
Organization	6 (5.66%)	14 (13.2%)	5 (5.88%)	37 (34.90%)	24 (22.64%)	5 (5.88%)
Content	4 (3.47%)	18 (15%)	3 (5.55%)	38 (33.04%)	32 (26.66%)	7 (12.96%)
Vocabulary	3 (5.08%)	16 (16.49%)	2 (5.71%)	36 (61.01%)	36 (37.11%)	4 (11.42%)
Total	28	76	15	184	137	26

*CGF= concise general feedback, EF= elaborated feedback, JEF= justified elaborated feedback

Reasons for Feedback Non-Incorporation

To collect the participants' reasons for not incorporating peer feedback on their master's thesis proposals, the researchers employed eye-tracking data to stimulate the participants to talk about their thoughts while applying peer comments on their texts. The thematic analysis of the interview data showed four main reasons for not applying the comments (Table 6).

Table 6

Extracted Categories of Reasons for Not Applying Peer Comments on Academic Texts

Theme	Sub-themes
Too general	- Not specifying the erroneous item among correct ones - Not providing steps to take to improve the text
Inapplicable	- Perceived low level of writing or subject-area knowledge - Unrealistic expectations of the feedback provider
Faulty	- Not being compatible with the instructor's guideline - Providing faulty information
Incomprehensible	- Including difficult lexical items or grammatical structures - Using symbols (such as question marks or exclamation marks) without any word - Using jargons

Table 7

Reasons for Not Applying the Comments

Reason	Frequency	Percentage
Too general	152	32.61
Inapplicable	60	12.87
Being faulty	125	26.82
Incomprehensible	129	27.68
Total	466	100

As reported in Tables 6 and 7, one of the reasons that the participants mentioned for not applying the comments was the feedback specificity level. The participants stated that comments which were too general were not easy to incorporate. Two sub-themes extracted from the interviews were identified in learners' responses. The participants had difficulty applying comments when the comment included a general requirement or asked for a modification without elaboration. The followings are examples of these sub-themes.

> *I did not apply this comment because I couldn't do so! It asked me to remove unnecessary 'the's from the text, but which ones were erroneous? Which ones were correct? So, I decided to ignore this comment.*
>
> *The comment asked me to work on the logical order of the arguments, but how could I do that? If the feedback provider were more detailed, maybe I could modify my text.*

The second reason reported by the participants addressed the inapplicability of some comments. The participants mentioned this factor 60 times (12.87 %) as the cause of their unincorporated comments. The following two quotations are examples of this theme.

> *I think a very good student had provided comments on my text because the comments were so difficult. I think my introduction was good, but he/she wanted me to write a more persuasive one to show the significance of the study. I believed that I could not apply this comment, so I did not apply it.*
>
> *I deleted these two sentences since I could not elaborate on the issue. I think this level of elaboration is for a Ph.D. dissertation and not an MA thesis, so I deleted the sentences not to be required to write supporting sentences which were beyond my ability.*

The third reason for not applying comments appeared to be that the participants' considered some comments as faulty. More than a quarter (*N* = 125, 26.82 %) of the unincorporated comments were reported to be faulty. The participants found these comments faulty or against what is suggested in well-known references or their instructors' instructions. Examine the following quotes taken from the stimulated recall interviews.

> *Why didn't I apply this comment? It was wrong. This comment asked me to write independent samples t-test to compare four groups. No, I had to write One-way ANOVA, and I didn't change it.*
>
> *Our instructor had asked us to avoid writing hypotheses when the study was descriptive, but this comment had asked me to write hypotheses. She/he has written in capital letters too (laughing)!*

The last reason for not applying comments was said to be comment incomprehensibility. The participants stated that some comments were difficult to understand; therefore, they did not/ could not apply them. Around a quarter of unincorporated comments fell into this category (*N* = 129, 27.68 %). The followings are some quotations from the interviews.

> *[Talking about a reduced conditional type 2] I did not understand this comment. I had some thoughts, but I decided to avoid taking risks, so I omit my sentence.*
>
> *[Talking about three question marks] I didn't understand this comment. What's wrong with my paragraph? I didn't understand its meaning, so I ignored it.*
>
> *I didn't understand the meaning of 'heritage learners', and I thought my text was OK without this phrase, so I ignored it.*

Discussion

The present study examined how Iranian L2 master's degree students engaged with peer feedback on their thesis proposals. The objectives of this study were to examine the extent to which master's degree students apply peer comments on their thesis proposals and why they tend to avoid some peer feedback on their thesis proposals. This study also studied how feedback types could affect the participants' accuracy of feedback incorporation. The results of this study supported

the feedback engagement model provided by Ellis (2010) by showing how different factors can result in learners' feedback use; however, the following paragraphs provide the discussion of results using prior theories and empirical studies.

The findings showed that L2 master's degree students incorporated around two-thirds of their peers' comments, which seems to be a high level of behavioral engagement. However, the analysis of the data indicated that feedback content in terms of specificity could noticeably affect graduatfication in addition to a detailed correction were the most successfuL2 learners are involved in the process of a more complicated task to identify the requirements of comments (Lachner & Neuburg, 2019). Furthermore, general comments can also impose adverse affective effects on learners, resulting in higher anxiety levels as they are less sure how to apply a comment which does not provide specific instruction on how to proceed (Fernando, 2020). This uncertainty can also result in negative feelings that can exacerbate the situation since the negative affective engagement with a comment can disrupt the cognitive and behavioral engagement of L2 writers. The resultant uncertainty may lead to learners' adversely-affected task self-confidence and motivation (Stevenson & Phakiti, 2019). The lowered self-confidence and motivation can, in turn, impact the behavioral and cognitive engagement with other comments in the same and subsequent feedback incorporation tasks.

The analysis of eye-tracking and stimulated recall interviews revealed four main reasons why the participants did not incorporate the comments: being too general, inapplicable, faulty, and incomprehensible. The effect of comment specificity on L2 writers' cognitive and behavioral engagement was discussed above. Master's degree students also stated that some comments were inapplicable. They argued that these comments were too difficult for them, that is beyond their abilities or skills, so they had no choice but to ignore them. Prior studies have also reported the adverse effects of the mismatch between learners' knowledge and the provided comments (Davin, 2013; Herazo et al., 2019). The suitability of comments for learners' (perceived) level has been controversial since the 1980s. Several scholars have discredited feedback activities as useless or even harmful if they are not matched with feedback receivers' knowledge (Truscott, 1996). Even in approaches, such as sociocultural theory, where feedback is a crucial factor in learning, the issue of reciprocity, which reflects how learners respond to mediation that has been

offered (Poehner & Wang, 2021), plays a significant role in the success of feedback activities. This responsiveness, which reflects learners' engagement with feedback, has been identified as a significant factor since comments do not modify learners' cognitive structures if learners have not reached the required cognitive ability. Moreover, it must be noted that there could have been some instances in which learners were behaviorally (as they applied the comments) but not cognitively engaged with comments (they did not learn them through cognitive/metacognitive strategies). Case studies can be conducted to identify the cases in which cognitive engagement is missing while behavioral engagement is evident.

The third issue that the participant mentioned for not applying the comments was the faultiness of the comments. The examination of peer feedback literature shows that there are reservations about the accuracy of peer feedback. Prior studies have shown that inaccurate peer comments can have adverse effects on learners' perceptions of peer feedback (Kaufman & Schunn, 2011; van der Kleij & Lipnevich, 2020), which can, in turn, decrease feedback receivers' engagement with the incoming comments since they are not sure if the comments include accurate information (Patchan & Schunn, 2015). Trust has been identified as a significant factor in L2 learners' feedback-seeking behaviors and can affect feedback recipients' engagement with comments (Sedikides et al., 2016). The literature on peer feedback displays that learners are less trustful when they feel the comments are provided by a less competent peer (Bahari & Gholami, 2022; Patchan & Schunn, 2015; Zhai & Ma, 2022). The extraction of this theme from the interview data shows that even in graduate writing contexts, the issue of trust is a significant factor that should be taken into consideration. Master's degree students with different levels of L2 writing ability and subject-area knowledge took part in this study, and the disparity between their knowledge might have resulted in the comments perceived as inaccurate. However, the participants' assessment of comments as inaccurate can be examined in another study.

The fourth factor mentioned as the reason for not incorporating comments was the incomprehensibility of comments. The first step of applying a comment is understanding its meaning, and applying an incomprehensible comment is unlikely to happen (Fan & Xu, 2020; Han, 2017; Patchan & Schunn, 2015). Previous studies have shown that one of the disadvantages of written comments is that feedback

providers cannot realize the extent to which their comments are understood until they read the revised version (Ellis, 2010), and feedback receivers are not able to ask for clarification immediately when feedback provider is not readily available to answer. This temporal gap can result in feedback receivers' inability to understand comments or misunderstanding them, which can disrupt the process of feedback incorporation.

Sasch and Polio (2007), emphasizing the significance of feedback understanding, argue that feedback receivers' identification and understanding should be ensured to make their feedback uptake possible. The examination of the dual-layered awareness in a feedback activity has shown that the noticing level of awareness is not sufficient, and students should reach the level of understanding to benefit from a comment (Rosa & Leow, 2004). In the same line, in his oft-cited model of second language acquisition, Robinson (1995) argues that although considerable currency has been given to alertness and orientation, the golden gate of learning an item is detection, which is the cognitive registration of the provided stimuli and can be actualized solely when a learner understands the provided stimuli. These insights from the literature vividly reveal the significance of understanding in learning. In the present study, it was witnessed that students' difficulties in understanding comments accounted for one-third of all ignored comments.

The findings also showed that peer feedback types could affect the participants' accuracy of feedback incorporation, and L2 master's degree students applied justified elaborated comments more accurately than concise general and elaborated comments. It seems that when feedback providers support their comments with justifications, feedback receivers have a less difficult cognitive task of incorporating comments into the revised version accurately since these justifications can guide feedback receivers on how to apply the comments (Gu'enette, 2007; Zhu & Carless, 2018). These justifications are reported to increase feedback receivers' engagement with comments and enable them to access their previously learned items more easily (Fernández-Michels & Fornons, 2021). This can increase the chances of high-quality revisions and learning (Berndt et al., 2018; Bolzer et al., 2015; Walker, 2015). In addition, justification can function against learners' negative perceptions of peer feedback credibility since feedback receivers are provided with explanations for the provided suggestion/correction (Han &

Hyland, 2019; Yu et al., 2019). This positive perception can, in turn, increase learners' attention to the feedback and increase the chances of accurate feedback incorporation.

Conclusion

Drawing on the data collected using eye-tracking technology, stimulated recall interviews, and content analysis, we examined L2 Iranian master's degree students' engagement with peer feedback on their academic texts. Based on the findings of this study, L2 graduate students apply justified elaborated feedback more than elaborated and concise general feedback. Similarly, the lowest level of ignoring comments belonged to justified elaborated feedback; text modifications to avoid feedback incorporation are not noticeably different across the three feedback types. In line with previous studies (Bai & Hu, 2016; Berndt et al., 2018; Mohammed & Al-Jaberi, 2021; Walker, 2015), this research shows that feedback type can project its effects on L2 postgraduate students' feedback incorporation decisions and those comments which include justifications are more successful in enabling feedback receivers to apply them in their revised versions. Given these findings, teachers are recommended to invite their students to provide detailed and justified feedback on their peers' texts. This can both deepen feedback providers' understanding of the issue (Walker, 2015) and, as we witnessed in this study, affect the feedback incorporation rate.

Finally, the results of this study demonstrate that L2 MA students apply peer comments more accurately when they are detailed and justified rather than concise and general. L2 learners might benefit from the cognitive facilitators (e.g., elaborated explanations) that guide learners through the cognitively-demanding journey of feedback incorporation and help them uncover how to apply their peers' comments. Feedback receivers' emotional engagement with feedback has been reported to affect their cognitive and behavioral engagement, and positive feelings have been found to positively affect learners' feedback incorporation and learning (Han & Hyland, 2019; Zhang & Hyland, 2018). Again, these findings imply that second language instructors had better encourage their students to provide elaborated feedback plus justification to increase the accuracy of feedback incorporation.

There were some limitations to this study that can be mentioned here. First, the researcher used eye-tracking technology to collect a part of the required data, but a few participants (not more than five students) did not follow the instructions in the practice phase (before starting the data collection stage), which took some extra time and might have negatively affected their revision process. In addition, this study benefited from remote eye-tracking technology, which can be practiced when the participants are not physically available, but using more precise eye-tracking equipment could possibly provide us with more detailed data. Furthermore, the focus of this study was on L2 academic writing, and the findings have to generalized cautiously to other writing settings since the context of feedback studies has been reported to affect learners' needs and wants (Stevenson & Phakiti, 2019).

While the present study fills a part of the gap in the literature, and these findings provide empirical evidence for the significant issue of peer feedback incorporation o by L2 MA students, further studies can be conducted to examine the reasons behind unincorporated comments in other contexts (e.g., general English courses and undergraduate courses, supervisor feedback). Furthermore, other researchers can examine the extent to which different feedback types result in learners' uptake. Using eye-tracking data and stimulated recall interviews, other researchers can uncover the extent to which different feedback types could result in learning in the short and long-run.

References

Amiryousefi, M. (2019). The incorporation of flipped learning into conventional classes to enhance EFL learners' L2 speaking, L2 listening, and engagement. *Innovation in Language Learning and Teaching*, *13*(2), 147-161. https://doi.org/10.1080/17501229.2017.1394307

Aubrey, S., King, J., & Almukhaild, H. (2020). Language learner engagement during speaking tasks: A longitudinal study. *RELC Journal*, *53*(3). https://doi.org/10.1177/0033688220945418

Bahari, A., & Gholami, L. (2022). Challenges and affordances of reading and writing development in technology-assisted language learning. *Interactive Learning Environments, 32*(1), 35-50. http://dx.doi.org/10.1080/10494820.2022.2065308

Bai, L., & Hu, G. (2016). In the face of fallible AWE feedback: How do students respond? *Educational Psychology, 37*(1), 67-81. https://doi.org/10.1080/01443410.2016.1223275

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. https://doi.org/10.1007/s10212-017-0343-z

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. https://doi.org/10.1111/jcal.12091

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education*, *31*(2), 219-233. https://doi.org/10.1080/03075070600572132

Cutumisu, M., Turgeon, K. L., Saiyera, T., Chuong, S., González Esparza, L. M., MacDonald, R., & Kokhan, V. (2019). Eye tracking the feedback assigned to undergraduate students in a digital assessment game. *Frontiers in psychology*, *10.* https://doi.org/10.3389/fpsyg.2019.01931

Davin, K. J. (2013). Integration of dynamic assessment and instructional conversations to promote development and improve assessment in the language classroom. *Language Teaching Research*, *17*(3), 303-322. https://doi.org/10.1177/1362168813482934

Donia, M. B., Mach, M., O'Neill, T. A., & Brutus, S. (2022). Student satisfaction with use of an online peer feedback system. *Assessment & Evaluation in Higher Education*, *47*(2), 269-283. https://doi.org/10.1080/02602938.2021.1912286

Ellis, R. (2010). A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition, 32*(2), 335-349. https://www.jstor.org/stable/44488131

Fan, Y., & Xu, J. (2020). Exploring student engagement with peer feedback on L2 writing. *Journal of Second Language Writing*, *50(3), 65-85.* https://doi.org/10.1016/j.jslw.2020.100775

Fernández-Michels, P., & Fornons, L. C. (2021). Learner Engagement with Corrective Feedback Using Think-Aloud Protocols. *JALT CALL Journal*, *17*(3), 203-232. https://doi.org/10.29140/jaltcall.v17n3.461

Fernando, W. (2020). Moodle quizzes and their usability for formative assessment of academic writing. *Assessing Writing*, *46*. https://doi.org/10.1016/j.asw.2020.100485

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. https://doi.org/10.1016/j.learninstruc.2009.08.007

Godfroid, A., & Hui, B. (2020). Five common pitfalls in eye-tracking research. *Second Language Research*, *36*(3), 277-305. https://doi.org/10.1177/0267658320921218

Gu´enette, D. (2007). Is feedback pedagogically correct?: Research design issues in studies of feedback on writing. *Journal of Second Language Writing, 16*(1), 40-53. https://doi.org/10.1016/j.jslw.2007.01.001

Han, Y. (2017). Mediating and being mediated: Learner beliefs and learner engagement with written corrective feedback. *System*, *69*, 133-142. https://doi.org/10.1016/j.system.2017.07.003

Han, Y., & Hyland, F. (2019). Academic emotions in written corrective feedback situations. *Journal of English for Academic Purposes*, *38*, 1-13. https://doi.org/10.1016/j.jeap.2018.12.003

Herazo, J. D., Davin, K. J., & Sagre, A. (2019). L2 dynamic assessment: An activity theory perspective. *The Modern Language Journal*, *103*(2), 443-458. https://doi.org/10.1111/modl.12559

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. *Journal of Language Research, 11*(33), 327-352. https://doi.org/10.22051/jlr.2018.19992.1536

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. https://doi.org/10.1007/s11251-010-9133-6

Lachner, A., & Neuburg, C. (2019). Learning by writing explanations: Computer-based feedback about the explanatory cohesion enhances students' transfer. *Instructional Science*, *47*(1), 19-37. https://doi.org/10.1007/s11251-018-9470-4

Mohammed, M. A. S., & AL-Jaberi, M. A. (2021). Google docs or Microsoft word? Master's students' engagement with instructor written feedback on academic writing in a

cross-cultural setting. *Computers and Composition*, *62*. https://doi.org/10.1016/j.compcom.2021.102672

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: perceptions of students and academics. *Assessment & Evaluation in Higher Education*, *42*(2), 266-288. https://doi.org/10.1080/02602938.2015.1103365

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. https://www.jstor.org/stable/43575308

Poehner, M. E., & Wang, Z. (2021). Dynamic assessment and second language development. *Language Teaching*, *54*(4), 472-490. https://doi.org/10.1017/S0261444820000555

Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *Journal of Second Language Writing*, *52* https://doi.org/10.1016/j.jslw.2021.100816

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, *45*(2), 283-331. https://doi.org/10.1111/j.1467-1770.1995.tb00441.x

Rosa, E. M., & Leow, R. P. (2004). Computerized task-based exposure, explicitness, type of feedback, and Spanish L2 development. *The Modern Language Journal*, *88*(2), 192-216. https://www.jstor.org/stable/3588751

Ryan, T., & Henderson, M. (2018). Feeling feedback: Students' emotional responses to educator feedback. *Assessment & Evaluation in Higher Education*, *43*(6), 880-892. https://doi.org/10.1080/02602938.2017.1416456

Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, *29*(1), 67-100. https://doi.org/10.1017/S0272263107070039

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment and Evaluation in Higher Education, 35*(5), 535-550. https://doi.org/10.1080/02602930903541015

Sedikides, C., Luke, M. A., & Hepper, E. G. (2016). Enhancing feedback and improving feedback: Subjective perceptions, psychological consequences, behavioral outcomes. *Journal of Applied Social Psychology*, *46*(12), 687-700. https://doi.org/10.1111/jasp.12407

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing*, *22*(3), 286-306. https://doi.org/10.1016/j.jslw.2013.03.011

Sinclair, H. K., & Cleland, J. A. (2007). Undergraduate medical students: Who seeks formative feedback? *Medical education*, *41*(6), 580-582. https://doi.org/10.1111/j.1365-2923.2007.02768.x

Stevenson, M. & Phakiti, A. (2019). Automated feedback and second language writing. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 125-142). Cambridge University Press.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*, 327-369. http://dx.doi.org/10.1111/j.1467-1770.1996.tb01238.x

Van der Kleij, F. M., & Lipnevich, A. A. (2021). Student perceptions of assessment feedback: A critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability*, *33*(2), 345-373. https://doi.org/10.1007/s11092-020-09331-x

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. https://doi.org/10.1080/02602938.2014.898737

Yu, S., Zhang, Y., Yao, Z., Yuan, K., & Zhang, L. (2019). Understanding student engagement with peer feedback on master's theses: A Macau study. *Assessment and Evaluation in Higher Education, 44*(1), 50-65. http://dx.doi.org/10.1080/02602938.2018.1467879

Yuan, J., & Kim, C. (2018). The effects of autonomy support on student engagement in peer assessment. *Educational Technology Research and Development*, *66*(1), 25-52. https://doi.org/10.1007/s11423-017-9538-x

Zhai, N., & Ma, X. (2022). Automated writing evaluation (AWE) feedback: A systematic investigation of college students' acceptance. *Computer Assisted Language Learning*, *35*(9), 2817-2842. https://doi.org/10.1080/09588221.2021.1897019

Zhan, Y., Wan, Z. H., & Sun, D. (2022). Online formative peer feedback in Chinese contexts at the tertiary Level: A critical review on its design, impacts and influencing factors. *Computers & Education*, *176*. https://doi.org/10.1016/j.compedu.2021.104341

Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, *36*, 90-102. https://doi.org/10.1016/j.asw.2018.02.004

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