Metacognitive Instruction for Peer Review Interaction in L2 Writing

Gavin Bui & Amy Kong

The Hang Seng University of Hong Kong | Hong Kong

Abstract: While previous research on peer feedback in L2 writing has stressed the importance of training in giving useful comments on peer's writing, very few studies have specifically explored metacognitive training in peer review interaction in terms of the perception and actual reviewing practices of L2 learners. This mixed-methods study employed a 12-week intervention course in L2 writing, in which eighteen Secondary One students, aged 12–13 years, received metacognitive training in peer review interaction. The training focused on metacognitive knowledge and metacognitive regulations. The results showed that, first, metacognitive training in peer review interaction helped change the perception of these young learners and increased their level of engagement and collaboration during the five peer review tasks. Second, metacognitive training appeared to encourage students to provide more content-related feedback than language-related feedback during dyadic interactions. Finally, it was found that the students were able to seriously consider their peer feedback when revising the peer-reviewed drafts, but written feedback had a much higher chance of being incorporated than did oral feedback. The pedagogical implications of these findings were discussed.

Keywords: metacognition, metacognitive training, peer review, L2 writing, peer review interaction

journal of WRITING RESEARCH

1. Introduction

The ability to write in English has been at the core of L2 teaching and learning worldwide. Writing requires sophisticated levels of cognitive functioning that may impose great challenges on all learners, especially younger ones. As a seemingly essential component of teaching L2 writing, the nature of and approaches to feedback have been controversial topics in the field, not to mention peer feedback (Lee, 2017; Ferris & Hedgcock, 2014). The effectiveness of peer feedback, in particular, has received sustained interest in the L2 writing literature (e.g., Hirose, 2012; Kong, 2019; Miao, Badger, & Zhen, 2006; Patchan & Schunn, 2016; Ruegg, 2015). Among the most commonly employed approaches to improving peer feedback in L2 writing are various types of "training" (e.g., Villamil & De Guerrero, 1996) or "modeling" (e.g., Ren & Hu, 2012). These interventions have been proven as useful cognitive strategies. On the other hand, no study appears to have explored the effects of metacognitive skills on the peer review process. Given the importance of metacognition in general learning (Flavell, 1976; Scott & Levy, 2013), it is surprising that the use of metacognitive strategies for peer review continues to be ignored in the L2 writing classroom. In view of the underexplored area of peer feedback among young learners, who have yet to develop the full potential of their cognitive and metacognitive skills, this study explored metacognitive training for peer feedback interaction in L2 writing, focusing on how metacognitive knowledge and regulation (Lee & Mak, 2018; Qin & Zhang, 2019; Van Velzen, 2016) shaped the peer feedback process and subsequent revisions among the secondary school students in Hong Kong.

This study probed into the interplay between metacognition and peer review in secondary school L2 writing. The following sections were explored: (1) definitions of metacognition and its components in relation to L2 writing, (2) training for peer feedback in L2 writing, (3) types of peer feedback, and (4) writer response and revision after peer review.

1.1 Metacognition and L2 writing

The concept of metacognition has been widely applied in the psychology of learning, even though demarcating its boundaries has proven challenging (Scott & Levy, 2013). Metacognition has been defined as high-order thinking skills that involve knowledge and control of the cognitive processes (Flavell, 1976). There is a plethora of literature revealing a close association between metacognitive awareness and learning effectiveness (e.g., Baker, 1989; Lee & Mak, 2018; Scott & Levy, 2013; Teng, 2017, 2019; Zhang, 2001, 2010). The researchers of L2 writing discussed components of metacognition similar to each other (Lee & Mak, 2018), including metacognitive knowledge, metacognitive skills/strategies, and metacognitive experiences (feelings and judgments) in L2 writing. Metacognitive knowledge and regulations have been recognized across studies on general metacognitive models (Bereiter & Scardamalia,

2013; Chamot, 2005; Schraw, 1998) and metacognitive theories (Van Velzen, 2016), as well as studies on peer review in L2 writing and metacognitive strategies (Liu & Lin, 2007), as the two most important components of metacognition. Van Velzen (2016) argued that metacognitive experiences seem to "resolve into an impression or tacit metacognitive knowledge that may not be useable to direct learning" (p.19). Therefore, metacognitive experiences were not taken into consideration in the metacognitive training performed in this study.

Metacognitive knowledge includes knowledge about one's cognitive processes, which contain *task*, *person*, and *strategy* (Flavell, 1979). According to Lee and Mak (2018), *task* knowledge involves the knowledge of the purpose, nature, and demands of learning tasks. *Person* knowledge involves one's own understanding and assessment of oneself, such as aptitude, motivation, and age, in addition to one's cognitive processes. Finally, *strategy* knowledge involves the knowledge of cognitive strategies that are facilitative in attaining learning goals.

Metacognitive regulations can be further classified as planning, monitoring, and evaluating. According to the new taxonomy of metacognition in L2 writing (Lee & Mak, 2018), planning refers to identifying and analyzing a problem by students, in order to solve it with a selected strategy. Monitoring is related to the evaluation of one's cognition and efforts toward a certain task. In slight contrast to the previous two subcomponents, evaluating generally happens after a learning task is completed, which evaluates the quality of the task. Hence, it is concluded that the new taxonomy of metacognition in L2 writing (planning, monitoring, and evaluating) is similar to the three necessary metacognition regulatory skills proposed by Schraw (1998) and Karlen (2017). Moreover, this new taxonomy is also similar to the kinds of higher mental processes of writing suggested by Bereiter and Scardamalia (2013), which include goal setting, planning, memory search, problem-solving, evaluation, and diagnosis. The ability of successfully implementing such regulative strategies helps in enhancing one's self-regulating learning ability in L2 writing (Teng & Huang, 2019; Zhang, 2010), hence provides better sustainability and self-efficacy in learning. However, Abba, Zhang, and Joshi's (2018) study revealed that students' acknowledgement of metacognitive knowledge did not significantly impact writing performance.

1.2   General and metacognitive training for peer review

The research on peer feedback emphasizes the importance of optimizing its benefits for developing the writing literacy of students. In order to achieve this objective, strategy training for maximizing the effectiveness of peer review has been recommended in the literature, including improving students' awareness about the benefits of the activity (Hansen & Liu, 2005), apprising them for effective strategies and appropriate language for implementing the activity (Rollinson, 2005), and modeling the activity process (Ren & Hu, 2012). In principle, these approaches are classified into the category of cognitive strategies; however, metacognitive strategies, which are high order thinking skills, have not been discussed in detail in the past studies. In the following section, we present

past literature on strategy training for peer feedback in L2 writing and attempt to identify research gaps in metacognitive training in peer review.

1.2.1 Cognitive strategies

A large number of studies have identified different peer scaffolding strategies in peer interactions. Stanley (1992), for instance, identified the following seven categories of peer scaffolding strategies for a reader to implement: pointing, advising, collaborating, announcing, reacting, eliciting, and questioning. In a similar manner, Zhu (1995) followed the framework of Stanley (1992) by excluding collaborating and adding elaborating, hedging, confirming, and justifying. Similarly, Villamil and De Guerrero (1996) also identified various scaffolding strategies that include advising, eliciting, reacting, requesting clarification, announcing, instructing, and giving directives. On the other hand, Liu and Hansen (2002) suggested only the following three types of effective comments: evaluation, suggestions, and asking for clarifications. The categories of the scaffolding strategies suggested in the previous studies (i.e., Liu & Hansen, 2002; Stanley, 1992; Villamil & De Guerrero, 1996; Zhu, 1995) could be reorganized as a set of simplified but more systematic cognitive strategies for peer review training. The first strategy is "advising," which is similar to the "suggestions" in Liu and Hansen (2002). As mentioned in all the previous studies, the second strategy is "eliciting", which is a strategy that could effectively induce the writer to think about and rectify a mistake. The third strategy is "reacting", which refers to the evaluation of remarks and further explanation of the remarks. The reacting strategy is similar to the "evaluation" of Liu and Hansen (2002). As discussed in Liu and Hansen (2002), the fourth strategy is "asking for clarification", which is similar to "requesting clarification" and "questioning" as mentioned in the previous studies. As discussed by Villamil and De Guerrero (1996), the final strategy is "instructing", which may be useful in peer interaction between high-proficiency learners and hence it should be introduced in peer review training.

Based on the previous studies, the suggested set of cognitive strategies is listed below with examples.

Advising: I think you can use more adjectives for describing your teachers.

Eliciting: Something is incorrect here. Which tense should you use?

Reacting: The details of your school facilities are not good enough.

Asking for Clarifications: I do not quite understand here. Why were you late?

Instructing: You only have one pet, so you should not use the plural form "pets."

It should be noted that, in the current study, as the peer review training is specifically targeted at instructing student reviewers how to effectively express comments to their peer, the follow-up responses (e.g., confirming and justifying) were

not covered in the current study. For detailed examination of the peer's responses to different strategies, please see Kong and Bui (2019).

1.2.2 Use of appropriate language and metalanguage

Hansen and Liu (2005), similar to Ren and Hu (2012), suggested it is important to create purposeful peer response sheets for learners to follow. Therefore, it is considered helpful for teachers to provide students with a checklist so that they have something to refer to in the case of helplessness. In addition, students, especially junior students, may lack language expressions to comment on the peer's writing, for instance, commenting on the climax, plot, and characterization when reviewing story writing. It is mandatory, for implementing the activities, to have a provision of prior teaching so that students may use appropriate language expressions (Bei, 2011; Bui, 2018; Bui & Yu, 2019) to comment on the peer's work.

1.2.3. Effectiveness of training on quality of feedback and revisions

In the previous studies, the effects of training on peer review practices have been a constant concern. Stanley (1992) advised his students to undergo a fairly lengthy coaching procedure, which involved role-playing and analyzing evaluation sessions, discovering rules for effective communication, and studying the genre of student writing. In order to evidence the effectiveness of the coaching, the subsequent revised drafts of peer review sessions were analyzed. The results, when compared with the uncoached students, showed that the students who received coaching demonstrated a greater level of engagement in peer review interaction and provided clearer guidelines for the revision of drafts. However, Stanley (1992) did not report whether the revision led to improvement in the writing quality. Berg (1999) conducted a similar study that examined the effects of trained peer response on the revision type and writing quality of ESL students. The results showed that the revised drafts of the students who received feedback from trained peer reviewers demonstrated greater improvement in their revised drafts and provided more meaningful revisions as compared to those who received feedback from the untrained peer, regardless of the proficiency level. Hence, the results indicated that appropriate training imparted to peer reviewers on how to deliver effective peer feedback can help the peer writers in producing more meaning-oriented revisions, which in turn may result in a better quality of the second draft. Berg's results were in conformance with those of Min (2006), whose empirical study examined the impact of training on the effectiveness of peer review and observed that more responses were shared by the trained peer review groups and that the number of revisions with enhanced quality was significantly higher as compared to the number of revisions before training. Both Berg (1999) and Min (2006) did introduce the concept of peer review training in their studies, but the focus was not on metacognition. A more recent study by Teng (2016) examined the impact of the metacognitive instruction on the writing quality and observed that the metacognitive instruction with the cooperative

learning mode condition outperformed the cooperative-learning-only condition and the control group. However, peer review was not a concern in his study. It appears that no study so far has specifically investigated how metacognitive training in peer review practice could make a difference in L2 writing. Table 1 presents a summary of these studies.

Table 1. Studies on improving the effectiveness of peer review *on quality of feedback and writing/revisions* by training

Study	Setting/Participants	Results
Stanley (1992)	US 27 ESL students	Trained students produced more varieties of response types and pointing, while advising and collaborating provided higher instances of revisions.
Zhu (1995)	US 11 ESL learners	Trained students for peer response led to significantly more and better-quality peer feedback compared with the control group that received no such training.
Berg (1999)	US 46 ESL students	Compared with the untrained group, the reviewers in the trained group provided more quality peer feedback to their peer, hence helping the writers generate more meaning changes and produce better-revised drafts.
Min (2006)	Taiwan, China 18 university EFL learners	Training can positively enhance the impact of peer review in terms of revisions and writing quality.
Rahimi (2013)	Iran 60 ESL learners	Trained students focused both on formal and global aspects whereas the untrained ones focused on the formal aspect only, resulting in significant improvement in the writing quality.
Teng (2016)	China 120 EFL university students	The Cooperative learning + Metacognitive instruction condition yielded the highest mean scores in writing and regulation of cognition, followed by the Cooperative learning condition and the control group.

1.3   Types of peer feedback in L2 writing

The nature of feedback is often categorized considering the taxonomy of revisions of Faigley and Witte (1981). This involves either a surface (local) change or meaning (global) change. The former involves changes that alter sentence structures or use of language (e.g., tense, punctuation, and spelling), whereas the latter changes the meaning of the text by introducing new ideas or deleting the old ones. These meaning/global changes further include macro and micro changes. The macro changes of the text are major changes that affect the overall essence of the text, whereas micro

changes only revise the information structure (e.g., adding more elaboration) without affecting the main ideas.

The peer feedback procedure has always been criticized for focusing only on surface/local changes and vague comments, especially when students have not received proper peer review training (Min, 2005; Rahimi, 2013; Stanley, 1992; Tsui & Ng, 2000). Caulk (1994) examined the usefulness of peer feedback and observed that 89% of the feedback comments were valid. However, his study did not provide detailed definition of "validity," nor did he compare the difference between surface and meaning changes. On the contrary, Miao, Badger, and Zhen (2006) and Paulas (1999) provided a simple model for the students in the form of a feedback sheet with guided questions. They compared the differences between surface changes and meaning changes provided by the peer. Moreover, both observed that students made more surface changes as compared to meaning changes. A more recent study conducted by Chang (2015) reported a more focused and detailed modeling method. She, while demonstrating a model essay to her students, focused on commenting on global issues such as content and organization as compared to local issues such as grammar and vocabulary. The student reviewers were instructed to read the essay written by the peers without implementing any corrections and their main responsibility was to help the peer improve the writing content and coherence. Consequently, such directive modeling created a significantly higher percentage of global comments and lower local comments. This result was similar to a study conducted by Rahimi (2013), in which the students provided more global comments after receiving intensive training on how to provide quality comments with a balanced focus on local and global issues. However, the different results of these studies (Chang, 2015; Miao, Badger, & Zhen, 2006; Paulas, 1999; Rahimi, 2013) could be attributed to different research designs (unfocused modeling versus focused modeling). Considering these results, prior research has predominantly focused on written peer feedback only; whether face-to-face oral feedback makes a difference has not been explored. Therefore, there is still a lack of a comparison between these two modes in administering peer feedback in one single study. In this study, we have attempted to address these issues.

Though most of the relevant studies exclusively focused on adult learners, two recent studies were conducted in primary and secondary school settings. Dixon and Hawe (2017) investigated peer feedback practices in a primary school in New Zealand but their focus was on the perspective of the writing ability of teachers rather than the young learners. Coincidentally, Saliu-Abdulahi, Hellekjær, and Hertzberg (2017) also examined peer feedback in a secondary school in Norway from the perspective of a teacher. It has been observed that there is a lack of systematic investigation on the effectiveness of peer feedback among the learners at younger ages. Moreover, whether findings from students at the tertiary level can be generalized to other groups of learners, such as secondary students, remain unknown. This study attempts to respond to this research gap.

1.4 Writer response and revision after peer review

Whether students show uptake and incorporate peer feedback in their revisions has rarely been explored in the L2 writing literature. The uptake of peer feedback is defined in this study in terms of the writer's response or the actual use of feedback provided by their peers. Paulas (1999) observed that students incorporated more changes in meaning from peer feedback without providing any possible explanation of the results. However, in the study of 12 university students by Miao et al. (2006), there was a similar percentage of successful surface change per applicable peer feedback point (71%) to meaning change (61%). According to the interview data, one of the reasons was that the student reviewers wrote their comments at home before verbally explaining their written comments, be they surface or meaning, in the peer review sessions. The findings of Miao et al. (2006) highlight the importance of orally explaining the written feedback. Therefore, in this study, both written peer review and oral peer review were considered (as discussed by Chang (2016), Miao et al. (2006), and Tsui & Ng (2000)).

1.5 Research gaps and research questions

The current study was inspired by a few gaps identified in the literature. First, it appears that, despite a large number of studies on peer review training, few studies have been conducted on the systematic training of metacognition specifically for peer review, still less on metacognitive training for peer review interactions. This study focuses on how metacognitive training influences peer review practice. Second, peer feedback has been considered as a general umbrella term in past studies. A few past studies addressed the role/effectiveness of oral and written feedback from the perspectives of both the reviewer and the writer. The objective of this study is to distinguish between these two modes of peer feedback resulting from metacognitive training. Finally, the peer feedback literature appears to favor the tertiary context, with other contexts, such as secondary schools where writing plays an important role in L2 teaching, which looks under-investigated. The current research explores how metacognitive training impacts peer feedback behaviors of secondary school students of Hong Kong and their acceptance of peer feedback in subsequent drafts. The following three research questions guided this study:

1. Do students perform peer review in accordance with the metacognitive peer review interaction training?

2. What are the differences in the nature of comments between oral feedback and written feedback provided by the reviewers after receiving the metacognitive training for peer review interaction?

3. To what extent do the writers incorporate the oral feedback and written feedback provided by the reviewers after receiving the metacognitively trained peer review?

2. Method

2.1 Participants

Eighteen Secondary One students aged between 12 and 13 years were selected as the voluntary participants of the study. They belonged to two band-one secondary schools in Hong Kong, where English was the medium of instruction (EMI). According to the Hong Kong education system, generally, band-one students are academically more advanced as compared with band-two and band-three students. They demonstrate higher English proficiency than those coming from lower-banding schools. These participants comprised 6 male and 12 female students who had been learning English for 9–10 years. They were estimated to have reached the A2–B1 levels of the Common European Framework of Reference for Language. The native language of all the participants was Cantonese. Only a few participants reported that they could speak one additional Chinese dialect, such as Hakka ($n = 2$) and Southern Min dialect ($n = 1$) at home, though most of the participants could also speak basic Mandarin Chinese. The researchers notified the students and their parents that the data collected from the peer review processes, stimulated recalls, interviews, and classroom observation would be kept confidential and the permission of students and their parents before the research was obtained.

2.2 Peer review training: components and procedures

The participants attended 12 sessions of an after-school writing enhancement course that was taught by one of the authors. The time period of each session was around 60–90 minutes. The objective of the course was twofold, first, to teach students the structure of and appropriate language for English essays, and second, to provide students with peer review training. To achieve this end, we provided students five topics of narrative writing (see Table 3) where different requirements on language and textual organization according to their genres (e.g., a personal letter vs. a news report) were introduced.

In the first lesson, the participants were individually interviewed, using the semi-structured format, about their past peer review experiences and perceptions of peer review, which helped the researchers and the students to understand themselves (*person* knowledge; see Table 2). Moreover, the researchers conducted a metacognitive training session for peer review interaction after the interview.

All peer review training components were carefully mapped against the metacognitive knowledge and regulation discussed in the literature review section (metacognitive experience does not appear to be "trainable" as discussed above) as elaborated in Table 2. Based on Lee and Mak (2018), the set of taxonomy of metacognition in L2 writing peer review is developed.

Table 2. Metacognitive training on peer review in this study

Dimensions of Metacognition	Components within each dimension	Specific training details
Metacognitive knowledge	Task	At the beginning of the writing course, the following was explained to the students: • The teaching schedule (five narratives of different topics in a 12-week course); • Procedure of the teaching in each session: Instruction on one topic => writing => peer review in the next lesson => revisions at home and submission in the next lesson; • Goal of peer review: At the beginning of the course and each time before peer review, the participants were reminded that the goal of peer review was to improve the writing task collaboratively; • The structural elements and the language features of the specific writing task: Before each peer review, the students were introduced appropriate metalanguage expressions on the structural elements and language features of the specific writing task (e.g., Adding "subheadings" for Personal Homepage and using the simple past tense for writing a News Story). The students were also introduced a list of common language expressions that could be used for effective peer review interactions (e.g., "Can you tell me why you write xxx here?")
	Person / role	They were instructed to play the role of a reviewer or a writer alternatively. The students were trained, • as reviewers, to show respect to and interest in the peer's work; • as writers, to stay open-minded to the peer's comments.
	Strategy	The students were asked to go through the writing three times. First, they should read through the text and comment on the content pursuant to the structural elements listed on the peer evaluation form. Then, they could read it once again in detail to indicate the mistakes related to the selected language items on the evaluation form. Finally, they could skim the work for the third time to grasp an overall impression of its organization and style. They were also encouraged to write down comments with reference to the criteria listed on the form. Such a selective focus helped ease their cognitive load. The students were shown examples of socio-cognitive

		strategies through the peer review guideline sheet, which could be employed during peer interaction, including advising, eliciting, asking for clarifications, reacting (giving evaluations), and instructing. The students were reminded to give constructive comments (e.g., *you should talk about why you like your schoolmates*) instead of the vague one (e.g., *your writing needs improvement*) to ensure their peer knows specifically which parts need improving. The students were asked to use English during the interaction; however, they could adopt code-mixing—using both Cantonese (L1) and English (L2)—in case of predicaments to express themselves in English. Note: the students were reminded again to employ these strategies before each peer review.
Metacognitive regulation	Planning	The peer evaluation form was distributed to the students each time before the peer review as a guide for planning the peer review process in the following areas: • content and structure • language use • genre and overall impression
	Monitoring	Students were trained to • go through the structural elements and language features listed on the form one by one; • put a tick in the boxes on the form or write down comments on the form after finishing an item; • ask their peer for clarification if they are not sure about the peer's work; • seek help from the teacher if they are not sure how to comment on the text.
	Evaluating (reflecting)	The students were invited to evaluate their own performance in the peer review in stimulated recalls one day after the peer review was conducted. They were guided to reflect on the reasons • why certain strategies were employed and • why certain review behaviors occurred. However, the researchers did not provide any evaluation.

Next, the second lesson began with the instruction on the first writing topic, which was then followed by students finishing the writing task at home. The students were then paired up and provided 30 minutes to execute peer review in the following lesson. A peer evaluation form was distributed to each student for reference before peer review.

The researcher also studied the structural elements and language features on the form in order to ensure they knew what and how to comment. The teacher also reminded them of the objective of the activity, the use of metacognitive knowledge and strategies, and appropriate expressions (as instructed in the formal training session in the first week) during peer review. In order to explore the reasons why they adopted certain strategies during peer review, stimulated recalls were conducted one day after peer review. The first, third, and last peer review processes and stimulated recalls were audio recorded for coding and analysis.

After the whole course, the students participated in semi-structured interviews conducted by the researchers about their perceptions of training and peer review. In order to facilitate their recalling process, all the interviews and stimulated recall sessions were conducted in Cantonese, which was their mother tongue. The sample pre/post-training interview and stimulated recall questions are presented in the Appendix. The detailed teaching and data collection procedure is presented in Table 3. In addition, one of the authors prepared classroom observation notes during all the training sessions.

Table 3. Teaching schedules (lesson plan, metacognitive training, and data collection)

Week	Activities	Remarks
Week 1	Two semi-structured interviews with all the students	Interviews
	Formal metacognitive training for peer review interaction	recorded
Week 2	Teaching of U.1 Personal Homepage (first draft)	
Week 3	Going through the peer evaluation form specific to U.1 and	Peer
	the gist of the earlier metacognitive training.	interactions
	Peer review of U.1	recorded
One day after Week 3	Stimulated Recall One with all students	Recalls recorded
Week 4	Collection of the first and revised drafts of U.1	
	Teaching of U.2 Diary Entry (first draft)	
Week 5	Going through the peer evaluation form specific to U.2 and	
	the gist of the earlier metacognitive training.	
	Peer review of U.2	
Week 6	Collection of the first and revised drafts of U.2	
	Teaching of U.3 News Report (first draft)	
Week 7	Going through the peer evaluation form specific to U.3 and	Peer
	the gist of the earlier metacognitive training.	interactions
	Peer review of U.3	recorded
One day after Week 7	Stimulated Recall Two with the selected students	Recalls recorded
Week 8	Collection of the first and revised drafts of U.3	
	Teaching of U.4 Personal Letter (first draft)	
Week 9	Going through the peer evaluation form specific to U.4 and	

	the gist of the earlier metacognitive training.	
	Peer review of U.4	
Week 10	Collection of the first and revised drafts of U.4	
	Teaching of U.5 Story-writing (first draft)	
Week 11	Going through the peer evaluation form specific to U.5 and	Peer
	the gist of the earlier metacognitive training.	interactions
	Peer review of U.5	recorded
One day after	Stimulating Recall Three with the selected students	Recalls
Week 11		recorded
Week 12	Final evaluation of the writing course	Interviews
	Two semi-structured interviews with all the students	recorded

2.3 Data coding and analysis

Two Ph.D. students in Applied Linguistics analyzed 30% of the data. They coded the first 5% of the data together, discussed the differences, and reached an agreement in all cases. Then, they coded the next 5% of the data. Thereafter, they independently coded 10% of the data and resolved all inconsistencies between the two coders through discussion in order to reach an agreement. They followed the same for another 10% of the data and observed that very little disagreement remained. Therefore, one of the coders analyzed the remaining 70% of the data. In the following sections, we explain the coding schemes that would enable systematic analyses of the data to answer the three RQs.

2.3.1 Types of episodes

The audio-recorded pair-talk sessions were transcribed and coded into different episodes, which were a segment of the pair talk during which learners explicitly focused on the narratives. Each episode was further coded as an on-task or an off-task episode (De Guerrero & Villamil, 1994). On-task episodes are utterances concerning the topics related to one discrete trouble source or a series of connected trouble sources, whereas off-task episodes are units of discourse in which the participants talked about issues or aspects of their lives unrelated to the content of the composition. Moreover, on-task episodes are further divided into three categories: reviewer/writer interactive, reviewer/writer noninteractive, and teacher-student interactive (see Table 4).

2.3.2. Types of oral feedback

The reviewer/writer episodes (the pair talk between the reviewer and the writer) of the first, third, and last peer review sessions were further coded as content-feedback-related episodes (CREs), language-feedback-related episodes (LREs), and unusable-feedback episodes. Similar to the LREs in the study of Swain and Lapkin (1998), the LREs in this study also focused on the segments in the dyadic interactions during which learners dealt with language items.

Table 4. Types of episodes (De Guerrero & Villamil, 1994)

On-task (Reviewer/Writer Interactive)	On-task (Reviewer/Writer Non-interactive)	On-task (Teacher-Students)	Off-task
Episodes in which both the reader and the writer discuss the revision of a trouble source.	Episodes in which only the reader or the writer talks about how to revise a trouble source without intervention from either party.	Episodes in which students talk to the teacher about some types of revision. The students may initiate the conversation after calling the teacher or the teacher may start the talk as she goes around.	A unit of a discourse in which the participants are talking about issues or aspects of their lives unrelated to the content of the composition.

However, the segments during which students explicitly focused on the content and structural development of the essays were defined as CREs. On the other hand, the CREs were further subdivided into macro-related and micro-related segments. According to the classification of the types of revisions by Faigley and Witte (1981), oral comments made to affect the development of the ideas of the whole essay were defined as macro-related segments. On the other hand, oral comments made to arouse simple adjustments to the existing text without affecting the overall flow of the essay were defined as micro-related episodes. Moreover, vague and implicit comments, such as "your story is quite boring" that cannot be successfully interpreted or used by the receiver, are referred to as unusable feedback (Hyland, 1998). Finally, we found the last category, erroneous-feedback episodes, which was not discussed in detail in the above-mentioned studies, emerged from the data but it did not fit into the existing coding schemes, so it was also included in this study. The recorded episodes that were coded into different types of feedback with episodes from this study are presented in Table 5.

Table 5. Different types of oral feedback

Type	Examples from the current study
CREs (Macro)	There are too many dialogues. I think the ratio of narration to dialogues is not balanced. Please reduce the number of dialogues.
CREs (Micro)	I think you can mention the woman's condition in the hospital here…
LREs	S1: My class teaches…something wrong here… S2: We have two teachers? S1: You have missed the letter "r" …
Unusable	Please write more examples to improve the essay quality. (It is unusable as it does not specify where to add the examples and what examples are needed.)
Erroneous (LREs)	I phoned Amy…Shall we add "to" here? So strange…

2.3.3. Types of written feedback

The first drafts of students were coded and analyzed to deal with different types of written feedback on the narratives of peers. Similar to that when coding different types of oral feedback, the written form of feedback was also categorized as content-related feedback that involves global comments made to the texts (macro and micro), language-related feedback, unusable feedback, and erroneous feedback. The language-related feedback, which dealt with only surface changes, was subdivided into additions, deletions, substitutions, permutations, distributions, and consolidations based on the study of Faigley and White (1981). Moreover, in this study, the students indicated errors by simply underlining or circling the trouble points. Therefore, these feedback points were considered as an indication in this study. The coding of the written feedback is presented in Table 6.

Table 6. Different types of written feedback

Code	Explanation	Example from this study
Content-related (Macro)	Make changes that affect the overall gist of the text	Maybe you can write about the other passengers' injuries (Please include the fact to what extent the passenger was injured)
Content-related (Micro)	Make changes that may only alter the information structure (e.g., adding more elaboration) without influencing the essence of the text.	What about the people (characters) of "One Piece" (comics)?
Language-related (Additions)	Add a word or phrase without changing the overall meaning of the sentence	I have black $^{\wedge straight}$ and curly hair
Language-related (Indications)	Indicate the trouble spot without changing anything	I like <u>play</u> with my pets.
Language-related (Deletions)	Omit a word or phrase without changing the overall meaning of the sentence	…. five people were hurt and sent to ~~the~~ hospital.
Language-related (Substitutions)	Exchange words with a synonym	Amy <u>go</u> to Disneyland _{went}
Language-related (Permutations)	Rearrangement of words or phrases	I like to play computer online games -> I like to play online computer games.
Language-related (Distributions)	One segment is divided into more than one	They are great parents, sometimes quite bossy. -> They are great parents. But sometimes they are quite bossy.
Language-related (Consolidations)	Combine two or more segments into one	Amy came. I was very amazed. -> Suddenly Amy came and I was

		very amazed.
Unusable	Vague and implicit comments that cannot be interpreted or used successfully by the writer	(0% in this study)
Erroneous	The comments are not correct.	He was talking on? this mobile phone.

2.3.4. Ratio of peer feedback incorporated into subsequent drafts

All the oral peer feedback provided in Tasks 1, 3, and 5 were recorded and transcribed. The transcription, together with the written peer feedback, was applied for analyzing the feedback that had been used in the subsequent drafts after metacognitive training in the peer feedback practice. In this case, a simple point-to-point comparison was made between the reviewer's comments and writer's responses. Moreover, the rejection of a comment (such as an erroneous one), as far as the writer justified it, was considered as feedback implementation. However, the ignorance of feedback or writer's self-revision was not considered as feedback implementation.

2.3.5. Coding of metacognitive knowledge and regulation

As shown in Table 2, the peer-to-peer interaction and interview transcripts were coded and analyzed. Moreover, we explored whether students understood the goal of the activity, divided their roles as a reviewer and a writer, and adopted the strategies and expressions introduced in the training in accordance with the components of metacognitive knowledge. Furthermore, we recorded whether students planned the peer review process by dividing it into three parts, namely, content and structure, language use, and genre. We employed a peer review form for facilitating the process and writing the comments, requested clarifications from their peer, and took the assistance of the teacher similar to the process of metacognitive regulation. All students underwent the "evaluation" process of metacognitive regulation as they were invited to retrospect their performance in the stimulated recalls.

On the basis of the metacognitive knowledge and metacognitive strategies used, five sociocognitive strategies were coded considering peer feedback interaction reported in previous studies (see Section 1.2.1), as presented in Table 7.

3. Results

The results pertaining to the three RQs are presented in this section. In-depth qualitative data, including stimulated recalls, post-course interviews, and textual analyses, are presented below. In order to triangulate the qualitative findings, quantitative data, where appropriate, were included.

Table 7. Coding of socio-cognitive strategies

	Description	Examples from the current study
Advising	Suggesting a revision or recommending the changes to be made	Maybe you can describe more about how you felt at that time?
Eliciting	Drawing out opinion, answer, or reaction	I wrote a letter…letter…. here…what's wrong?
Reacting	Making evaluative comments	I think you have included too many dialogues in the story.
Requesting clarification	Asking the writer to clarify or justify the intended meaning	Sorry…why do you say that your shirts are dirty?
Instructing	Giving mini-lessons on vocabulary, grammar, or other aspects of writing	Here…you have two people, so you should use the plural "are"

3.1 Do students perform peer review in accordance with the metacognitive peer review interaction training? (RQ1)

It was observed that training in metacognitive knowledge (Section 3.1.1) and metacognitive regulations (Section 3.1.2) changed the perceptions of students in their peer review process, and consequently, the behaviors of the students.

3.1.1 Metacognitive knowledge

Metacognitive knowledge consists of the knowledge of task, person, and strategy as discussed in Section 2.1. Different aspects of metacognitive knowledge in peer review practice are explored in the following with a focus on the knowledge of metacognitive strategies.

Metacognitive knowledge (task and person). The training session started with an overview of the objective of peer review, that is, how to collaborate with each other in order to improve the writing task. The students collaborated as a reviewer and a writer alternatively. Moreover, the students took their role very seriously during the peer review. For instance, all reviewers spotted the errors in the writing and explained them to the writer. They also recommended suggestions to the writer. In addition, they requested the writer to provide clarifications before making any changes when they were not sure about the ideas of writers. For instance,

Wendy: And here I don't know what is PS?

Hadley: Primary school…. sorry I wrote it too fast…

Note: Names of the participants in this study were all pseudonyms

Most importantly, the students were reminded to respect and appreciate the work of writers. It was observed that this kind of a positive attitude was evident in the actual peer review sessions wherein the reviewer complimented the writer. For instance,

Wendy: I think your language is pretty good

Hadley: I think overall, your essay is pretty good.

On the other hand, the writers were requested to be open-minded and welcome reviewer comments because the primary objective of the peer review was to improve the quality of the writing task. Moreover, all students were receptive to the reviewer's feedback during peer review. They carefully listened to their peer's suggestions. Considering the suggestions of the reviewers, all students reported making revisions in the final interview.

Most importantly, most writers followed the instructions given in the training and accordingly provided counter suggestions genuinely if they had a different opinion from the reviewer. In the following episode, Wendy believed that Hadley should have used "fun" instead of "funny" while describing her parents. On the other hand, Hadley believed that "funny" was the correct word because her parents always made jokes. Hadley humbly shared her justifications and they appeared reasonable. Wendy was a bit uncertain about her comment so took the help of the teacher.

Wendy: here… you should say 'fun' as you are talking about your parents, so you shouldn't use 'funny'?

Hadley: But I am saying they tell jokes, so I want to use 'funny'.

Wendy: Ummm… (Looking at the teacher)

Teacher: Maybe you can say 'My parents are fun because they always tell funny jokes.'

Hadley: Or…

To sum up, metacognitive training is considered helpful in terms of the knowledge of task and person for improving the secondary school students' awareness of the peer review processes and their appropriate roles as writers/reviewers.

Metacognitive knowledge (strategy): covering both global and surface aspects. As mentioned about the details of the training in Table 2, students were required to go through a narrative three times focusing on different aspects of writing, namely, content, language, and overall organization, at a time. All the students stringently followed the instruction. They enjoyed sharing comments on both language and content. For instance, in the first stimulated recall, Ophelia and Wendy stated that they followed the suggestion shared during training, such as reading the narratives two or three times in order to cover all aspects of the written work.

Ophelia: I tried to be a professional as I commented on both the content and language. I think both of them are important.

> Wendy: I followed the guidelines. First, I read the content. Then, I spotted the grammatical mistakes. After that, I talked to the writer about my comments. And finally, I ticked the form.

The quantitative findings supported these observations. In this study, students delivered language-feedback-related episodes most (51 episodes or 32.48%), followed by the macro-CREs (47 episodes or 29.94%) and micro-CREs (38 episodes or 24.20%). Moreover, there were more CREs (85 episodes or 54.14%) as compared to LREs (51 episodes or 32.48%) when the macro- and the micro feedback were combined (as total CREs). Comparatively, students shared little unusable feedback and erroneous feedback (17 and 4 episodes, or 10.83% and 2.55%, respectively). A Chi-square test showed that students produced significantly higher usable feedback as compared to the unusable ones or the erroneous ones ($p < 0.01$) in this study. Furthermore, students also significantly shared more feedback based on content as compared to language ($p = 0.03$). In other words, students were not only capable of producing valid and constructive feedback to their peers verbally, but this feedback also covered the surface as well as the meaning changes. These results contradict the previous research on the effectiveness of peer feedback, which is further elaborated in the Discussion Section.

In fact, in the first stimulated recall all students revealed that they liked the ideas of receiving/sharing different comments with regard to both the content and language of the narrative, such as the development and the creativity of the ideas as well as the language used in the narrative, which includes grammar, vocabulary, and spelling. Though some of them preferred one type of comment to another (e.g., content feedback over language feedback), they put their best efforts in order to share both types of feedback and expected to receive both types of feedback from their partner as well. For instance, Hadley preferred to share comments on content as she found it more interesting, whereas Wendy preferred to comments on language because she believed identifying grammatical errors was an easier task as compared to providing feedback on the content and organization. However, they both suggested that after metacognitive training, sharing feedback on both the content and language is equally important. Moreover, they also reported that it was useful to receive different aspects of feedback from their peers. For instance, Ophelia acknowledged the significance of both types of comments. Therefore, she attempted to share both types of comments in as much detail as possible. As she shared comments on both content and language during peer review, this activity shaped her role (as part of the metacognitive knowledge) as a more seasoned reviewer and reflected her positive attitude toward peer review.

Nonetheless, Ophelia had a few complaints about the comments shared by her partner, Alex. She reported that Alex failed to make any global changes to her writing, which was further confirmed by the researchers. During the interview, when asked about his preference for grammar feedback over content feedback Alex reported that he preferred to give language-related comments to Ophelia because he believed that the grammatical errors in her writing were not of very serious nature and she could still

manage to revise the paper according to his feedback. However, he realized that Ophelia's ideas were, to him, "good enough," which led to his lack of global comments on the work of Ophelia. This peer review process rendered Ophelia less satisfied with Alex's feedback, and she reported that her partner did not share in-depth comments on her works:

> Ophelia: His comments were quite superficial.

However, despite her dissatisfaction with Alex's imbalanced feedback, Ophelia still accepted his surface comments and made revisions accordingly. Moreover, she reviewed her essay once again in terms of the content and made revisions on her own in order to follow the metacognitive training.

In the final interview, when instructed to comment on the use of metacognitive strategies in peer review provided in the training, all students demonstrated a highly positive attitude toward such training because it provided clear directions to specific areas of concerns in peer review. The students reported that the guidelines of review strategies enabled them to prepare a schematic strategy when they planned their review process and evaluate the global and surface aspects of the narratives, for both peers and their own.

Metacognitive knowledge (strategy): use of sociocognitive strategies. As discussed in the literature review section, five sociocognitive strategies (advising, reacting, instructing, asking for clarifications, and eliciting) were introduced to students during the training (Villamil & De Guerrero, 1996). The students believed that these strategies could facilitate their peer review practice as revealed by the first stimulated recalls. These strategies did not only benefit their own reading of other students' works, but also allowed them to better understand the comments shared by their peers. Ophelia, Hadley, and Wendy reported that they received a deeper insight through interactions with their peers' feedback. Moreover, they also reported that training helped them to gain a better idea of what and how to interact. As reported by Hadley in the final interview:

> I love talking to my peer because it will be more impressive and detailed to know what she thinks about. It will be useless if we are advised to write on the form only.

The interview results were triangulated with the quantitative data (Table 8). Eighteen participants produced a total of 215 episodes on the application of the sociocognitive strategy in peer feedback interaction for the first, third, and fifth narratives, which showed that students followed these strategies frequently, with *advising* and *asking for clarifications* showing the highest frequencies. However, it was observed that students did not prefer *eliciting*. One of the students, Ophelia, explained why she preferred advising to other strategies:

> It was only natural that I recommended a suggestion whenever I spotted an issue. Other strategies might not come up in my mind at that time.

Table 8. Frequencies and percentages of different socio-cognitive strategies

	Advising	Eliciting	Reacting	Requesting Clarifications	Instructing	Total Reviewer/ Writer Episodes (Interactive + Non-interactive)
Homepage	32 (43.8%)	0 (0%)	14 (19.2%)	18 (24.7%)	9 (12.3%)	73
News	21 (34.4%)	0 (0%)	7 (11.5%)	24 (39.3%)	9 (14.8%)	61
Story	31 (38.3%)	1 (1.23%)	12 (14.8%)	25 (30.9%)	12 (14.8%)	81
Total	84 (39.1%)	1 (0.46%)	33 (15.3%)	67 (31.2%)	30 (14.0%)	215

Metacognitive knowledge (strategy): concrete comments with appropriate expressions. Moreover, students were reminded, in their role as a reviewer, to share valid comments instead of vague ones according to the planning component in the metacognitive strategies. Comments such as "your writing needs improvement" are considered vague and unusable because these do not indicate what and how the students should improve their writing. On the other hand, comments like "you should specify why you like your friends" are more valid because the peers know the specific areas that need reworking.

In this study, almost all the students demonstrated an ability to provide constructive and detailed comments in line with the objectives of the metacognitive training. For example, during the peer review of the homepage, they all indicated that on which particular sub-sessions (e.g., Me, My family, My hobbies, My school, My friends) their peers need to work on. Moreover, they also commented on the preciseness of paragraphs of the main body during the peer review of news reports and storytelling. For instance, Ophelia shared constructive feedback on how detailed the description should be:

How serious was the car accident? You should describe it in detail...

Hadley also specifically pointed out problems related to metalanguage:

Could you please use more adjectives to describe yourself in this paragraph?

Finally, the percentage of unusable feedback is quite low (10.83%), which supports the above self-reports of students.

3.1.2 Metacognitive regulations

In this section, how metacognitive training shaped perceptions of the participants and resulting behaviors in metacognitive regulations during peer review are discussed. Specifically, the understanding and actual implementation of strategies in planning, monitoring, and evaluating the review processes are described.

Planning: use of course materials, the training guidelines, and the evaluation form. It was observed that students referred to the training guidelines, peer evaluation form, and course materials quite frequently. These reference materials facilitated the interaction as students could always turn to the procedures, expressions, and vocabulary during peer review. In the last stimulated recall, all students reported that they followed the training guidelines and read the narratives based on the items listed in the evaluation sheet. For instance, Hadley knew her plan well before a review task:

> As instructed, I had to read my classmate's essay three times. I had to focus on language, content, and finally the overall structure each time.

An analysis of the interview transcripts showed that all students attempted to implement the metacognitive strategies although they varied in the extent to which they successfully followed the guidelines. Therefore, this may be considered as the result of their awareness of planning and task readiness (Bui, 2019; Bui & Huang, 2018) prior to peer feedback.

Monitoring. It was observed that students conducted peer review according to the guidelines mentioned in the evaluation form. Moreover, students made a final evaluation by marking on the form after exchanging their views with each other. They requested the writers for clarifications when they were not sure what was written. For instance,

> Wendy: … why were you late? (asking for clarifications)

> Hadley: My alarm did not ring. I wrote it in the next paragraph (offering justifications)

Furthermore, the students were also encouraged to consult the teacher in addition to requesting the writer for clarifications. In the first stimulated recall, all students reported that it was very useful to have external resources, such as feedback forms and the teacher, during peer review because these resources could facilitate their communication in English more effectively. They took the help of the external resources for contextually appropriate lexis and expressions during the course.

However, students shared few comments on the form after their face-to-face discussion. When questioned about their reluctance on providing written comments in the last stimulated recall, the students reported that writing feedback on a paper was more difficult and time-consuming as compared to directly discussing with their peers. Moreover, they suggested that direct conversation with peers was far more efficient as compared to written comments. As the training emphasized on oral interaction in peer

feedback, writing comments was only considered as a peripheral activity. As a result, in this course, the students achieved a collaborative decision primarily through verbal conversation but not written comments. Hence, this observation was found to be in line with the quantitative analysis of written feedback of students in the three written tasks (the first, third, and fifth narratives) as shown in Table 9 (the percentages are calculated considering all types of comments).

Table 9 presents the frequencies and percentages of the different types of written feedback. As observed in Table 9, relatively less content-related feedback (Macro + Micro: 16.14%) was provided by the students on the evaluation form in comparison to the language-related feedback (73.39%). Among the surface changes, substitution (23.39%), indications (21.56%), and additions (20.64%) were the three most frequently employed strategies. This finding implied that in written comments, the students were inclined to provide surface feedback by locating the errors directly in the peers' writing, and that these surface changes usually involved simple substitutions/additions of expressions or indications of the trouble spots. On the other hand, the students were inclined not to write any comments related to global changes in the narratives on the evaluation form. Moreover, in the written form, the students did not produce any unusable feedback. However, they did produce a considerably high percentage (11.47%) of erroneous feedback overall.

The results of the Chi-squared tests (Table 10) revealed that the students produced significantly more language-related written feedback than the content-related written feedback ($p < 0.01$). Moreover, no significant difference was observed between the content-related written feedback and the erroneous written feedback, implying that the students produced erroneous feedback at a relatively high rate.

3.2 Differences between oral feedback and written feedback after metacognitive training in peer review interaction (RQ2)

When analyzing content-related feedback and language-related feedback in detail, it was observed that the ratio of content-related feedback to language-related feedback was different between oral feedback and written feedback. Table 11 presents the different types of feedback produced by the students in oral and written forms. Numerically, the students produced higher oral feedback compared to written feedback in the case of meaning changes, while more written feedback was produced in comparison to oral feedback in the case of surface changes. On the other hand, the students produced a higher percentage of unusable oral feedback in comparison to the percentage of unusable written feedback, while the percentage of erroneous feedback observed in the written form was higher than that in the verbal form. Inferential statistics appears less appropriate in this case, as these are not strict comparisons since the students could choose whether to provide any written feedback or not.

Table 9. Frequencies and percentages of different types of written feedback

Content (33)		Language (160)							Unusable	Erroneous
16.14%		73.39%								
Macro	Micro	Addition	Indication	Deletion	Substitution	Permutation	Distribution	Consolidation		
(11)	(22)	(45)	(47)	(6)	(51)	(4)	(5)	(2)	(0)	(25)
5.05%	11.09%	20.64%	21.56%	2.75%	23.39%	1.83%	2.29%	0.91%	0%	11.47%

Table 10. Mean differences in the types of written feedback per student

Content = 0.58 (1.08)	Language = 2.81 (3.13)	$p = 0.00$
	Unusable = 0.00 (0.00)	NA
	Erroneous = 0.44 (0.85)	$p = 0.41$

Note. SDs in parentheses.

Table 11. Oral feedback vs. written feedback

	Content-related	Language-related	Unusable	Erroneous
Oral	54.14% (85)	32.48% (51)	10.83% (17)	2.55% (4)
Written	15.14% (33)	73.40% (160)	0% (0)	11.47% (25)
Total	31.46% (118)	56.27% (211)	4.53% (17)	7.73% (29)

The stimulated recalls data supported the quantitative findings. In the second stimulated recall, Hadley, despite acknowledging the importance of writing down the comments, confessed that she was not able to manage simultaneously writing down and verbalization of the comments. She said:

> I agree *[with Wendy]*. I also think that we should write down the comments. However, I want to express my opinions regarding her *[Wendy's]* ideas immediately as they arise in my mind during the peer review, and I probably won't be able to manage writing down the comments during the peer interaction as I would require a lot of time to think about how to write these ideas down."

In the second stimulated recall, Ophelia and Aaron stated that they read the writing first, had interactions, and marked off points in the evaluation form at the end of the peer review. The reason for following this sequence of steps was that both of them did not want to miss anything they intended to say during the peer interaction.

> Ophelia: "I read the writing, then talked to him *[Aaron]*, and marked off the points in the form at the end. I did not want to miss anything that I wanted to say, so I had to verbalize it all immediately. *[I do not need to write down or tick the points in the form during the peer interaction]* That way, he would understand all of it better."

In the final interview, when the students were asked the reason for their inclination toward providing more feedback related to global changes compared to surface changes during the peer interactions, all the reviewers stated that after training, they had a better understanding of the manner in which to interact with the peer, and therefore, it was more convenient and comprehensible for them to explain the problem with the content and organization development in their peers' papers during the peer interactions. For instance, Alex stated:

> I love your training because it taught me how to express my comments in (spoken) English.

On the other hand, the reviewers reckoned it to be extremely complicated for them to write down the comments associated with meaning changes. As a consequence, they preferred to express the sophisticated ideas associated with global changes verbally. For instance, in the final interview, both Hadley and Ophelia indicated a similar preference.

> **Hadley:** It's more comprehensible to talk to my peer for explaining certain difficult concepts rather than writing them down in the form.

However, in case of surface changes, it was observed that all the reviewers were inclined to simultaneously locate the grammatical mistakes directly while they were reading their peer's writing. In addition, as observed during the coding of peer review

episodes and students' writing drafts, most of the unusable feedbacks related to general comments on the narratives were produced orally. Unusable feedback was rarely observed in written form as the students were used to uttering such general impression verbally as their first instinct. On the other hand, the most erroneous feedback in the written form was associated with surface changes. Since the participants of the present study were a group of young L2 learners, they were not capable of expressing sophisticated ideas in completely accurate English. For instance, Aaron stated that he felt obligated to speak in English in class despite his reluctance to do so. Therefore, in order to facilitate communication, he opted for code-mixing.

> Aaron: I hate talking in English. However, since it's an English class, I attempt to converse in English. However, I think it is difficult for me to communicate in English. If I am not able to recall certain words, I return to speaking in Cantonese. I think it's a better idea to allow the students to switch between Chinese and English.

When compared with oral feedback, the students had less time pressure in editing the grammatical mistakes in the written form. Therefore, they provided more grammar-related feedback. At the same time, the chances of producing erroneous feedback were also higher in the written form of the feedback in comparison to the verbal one. However, it is noteworthy that irrespective of the nature of the feedback, the overall percentage of erroneous feedback remained low (7.73%) in comparison to the percentage of correct usable feedback (87.73%). Similarly, the overall percentage of usable feedback (87.73%) was significantly higher than that of the overall unusable feedback (4.53%) (Please refer to Table 11 for details).

3.3 Writer's incorporation of peer comments after metacognitive training in peer feedback interaction (RQ3)

This section reports the writer's incorporation rate of the written peer feedback, followed by their incorporation rate of oral peer feedback, in their subsequent revised drafts.

3.3.1 Incorporation of reviewer's written feedback in revisions

As observed in Table 12, the students exhibited ideally high percentages of incorporation of their peer's written feedback in their revised drafts of the first, third, and fifth essays. At least, 50% of any type of written feedback was incorporated into the revised drafts by the students. The type of written feedback with the highest percentage of incorporation among the students was deletion and distribution (100%), followed by substitution (88.2%) and addition (80%). The incorporation of the written feedback that involved meaning changes was also considerably high (Macro: 63.63%; Micro: 72.73%). These observations demonstrated the usefulness of written feedback.

Table 12. The percentages of the incorporation of different types of written feedback

	Type of feedback	Counts
Content-related Episodes	Macro	7/11 = 63.63%
	Micro	16/22 = 72.73%
Language-related episodes	Addition	36/45 = 80.00%
	Indication	34/47 = 72.34%
	Deletion	6/6 = 100.00%
	Substitution	45/51 = 88.24%
	Permutation	3/4 = 75.00%
	Distribution	5/5 = 100.00%
	Consolidation	1/2 = 50.00%

3.3.2 Incorporation of reviewer's oral feedback in revisions

Unlike the written feedback, for which generally higher percentages (>50%) of uptake were exhibited by the writers, varying percentages of incorporation of different types of oral feedback were observed in the revised drafts of the writers. Table 13 summarizes the general pattern of students' incorporation of their peers' feedback in their revised drafts. Overall, written feedback appeared to have a better chance of being incorporated in the later drafts compared to the oral feedback. In addition, language-related feedback was observed to be better incorporated in the revised drafts, irrespective of the form of the feedback (oral or written). While the rate of incorporation of the language-related oral feedback was not low (76.47%), students exhibited a relatively low percentage of the incorporation of the meaning-related feedback.

Table 13. Incorporation rate of the different types of oral and written feedback

	Content-related (Macro + Micro)	Language-related (Total)
Oral	Macro: 18/47=38.30%	39/51=76.47%
	Micro: 16/38=42.11%	
	Total: 34/85=40.00%	
Written	Macro: 7/11=63.64%	130/160=81.25%
	Micro: 16/22=72.73%	
	Total: 23/33=69.70%	

As inferred from the data obtained in the final interview, the main reason for the different uptake rates between the oral and the written feedback was that the students had forgotten what their peers had commented verbally during the peer review, especially when it came to complicated changes associated with global development. For instance, in the second stimulated recall, Wendy mentioned:

> I think verbal comments are good because I was able to understand her *[Hadley]* better. However, I hope she could also write down the comments as later I would probably forget what she had commented.

On the other hand, the reason underlying the higher incorporation rate of language-related feedback among the students was that most of the time, when the feedback was related to surface changes, the reviewers also marked their verbal feedback on their peers' narratives. In other words, the high percentage observed for the language-related changes in the revised drafts of students could have been affected by both oral feedback and written feedback.

4. Discussion and Concluding Remarks

The present study, which was designed as mixed-methods research, explored the impacts of metacognitive training in L2 peer feedback interaction on young learners. Given the non-experimental nature of the present study, the focus was placed on the manner in which students perceived the metacognitive training provided to them, and whether their peer review practices after the training were consistent with the goals of the training provided. While according to the literature, training in peer review in L2 writing was proven to be effective (Table 1), little research involved metacognitive training for peer review, let alone the training meant specifically for peer review interactions. The present study provided the Hong Kong secondary school students with training sessions that were carefully designed with a focus on their metacognitive awareness during the peer review practices (see Table 2 for training details). The findings of the present study (Section 3.1) suggested that students could enhance their metacognitive knowledge and metacognitive regulations from such a training, and the results appeared to counteract certain unfavorable results regarding the effectiveness of peer feedback reported in previous studies, including students' negative attitude toward the beneficial role of peer feedback (Amores, 1997; Nelson & Carson, 1998; Sengupta, 1998; Zhang, 1995), and the student reviewers' preference for surface changes rather than global changes during the peer review (Caulk, 1994; Miao et al., 2006; Paulas, 1999). The post-course interviews revealed that the students welcomed the opportunity of training regarding what and how to comment, and received guidance on planning the review process and monitoring their own progress using a checklist, while the stimulated recalls assisted them in reflecting on and evaluating their accomplished reviews, which could be incorporated into their next written assignment and the corresponding review session. In addition, it appeared that metacognitive training in peer review should not be restricted to the so-called cognitively "mature" students at the tertiary level; the training could also be applied to young learners such as those who participated in the present study.

Both qualitative and quantitative data revealed that the participants were inclined to provide global comments verbally, while the local comments were provided more in the written form. This finding provides new insights in comparison to those provided by the existing literature in the field (Paulas, 1999; Miao et al., 2006). The reasons for this difference in the nature of comments could be that the students found it more convenient to talk to their peers regarding sophisticated ideas rather than writing them down. This may also be related to the proficiency level of the students. For example, Alex reported that he loved the idea of speaking in L1 because he could opt for his mother tongue whenever he found it difficult to express himself in the English language. As described in Section 4.2, Aaron directly expressed his unwillingness to use L2 English, given his difficulty in expressing his thoughts in English; he instead preferred the Cantonese language for providing feedback. In the case of surface changes, the quantitative data revealed that "deletion" and "addition" registered the highest percentages. In other words, these Secondary One students could only manage to perform simple surface changes on the paper by deleting certain words or adding a few simple ones.

The relatively high rate of successful revision (•50% for written feedback and around 40% for oral feedback) indicated that the young learners participating in the present study could successfully incorporate their reviewers' feedback in their subsequent revision. However, the participants exhibited a higher incorporation rate of their reviewer's written feedback, rather than the oral feedback, in their revised drafts, especially in case of meaning-related feedback. This finding corroborated the findings of previous studies (Chang, 2016; Miao et al., 2006; Tsui & Ng, 2000) which stated that a combination of written and oral feedback appears to be the best compromising approach to enhance the effectiveness of peer feedback. However, this finding was in contrast with what students expressed in their final interviews. Most of the students preferred providing comments on global changes verbally, because of the preciseness and convenience of expressing their opinions, although the low incorporation rate of oral meaning-related feedback implied that the students had forgotten the provided feedback, and were not able to revise the drafts accordingly if feedback was not written down. Although the students exhibited a desirably high percentage of incorporating oral surface-changes feedback in their draft revisions, this could have happened because such surface changes were marked in the narratives prior to a conversation with the peers. In other words, although the students claimed that it was more comprehensible and convenient for them to express themselves by talking to their peers, and that in this way they were able to understand the reason for being asked to correct their writing in certain ways more conveniently, the most effective way for the students to "remember" incorporating their peer's feedback in their revised drafts after the class was still the written feedback. However, writing down their comments, especially those involving meaning-related changes, required high L2 proficiency and great knowledge in meta-language of the students. For instance, the students were expected to write down expressions such as "the ratio of narration to dialogues is not

balanced enough" on the writing draft or on the peer evaluation form. Such a requirement might have imposed a great challenge to the Secondary One students, even if they were previously asked to write this down in their L1. In the final interview, students stated that they were reluctant to write down their comments, especially when the comments were related to idea development because they considered it complicated and disruptive to the flow of interactions. As a result, the students preferred verbally discussing the ideas with their peers. Figure 1 summarizes the general findings of this research, along with relevant pedagogical implications that will be further elaborated in the sections ahead.
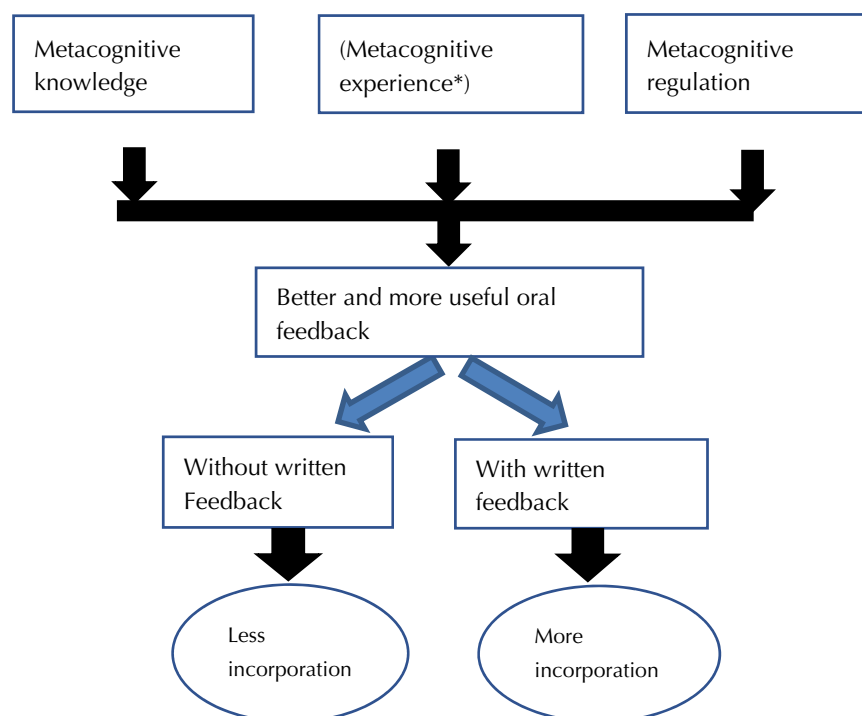
```
┌──────────────┐   ┌──────────────┐   ┌──────────────┐
│ Metacognitive│   │(Metacognitive│   │ Metacognitive│
│  knowledge   │   │ experience*) │   │  regulation  │
└──────────────┘   └──────────────┘   └──────────────┘
        │                  │                  │
        ▼                  ▼                  ▼
        ┌──────────────────────────────────────┐
                           │
                           ▼
              ┌───────────────────────────┐
              │ Better and more useful oral│
              │        feedback           │
              └───────────────────────────┘
                   ↙              ↘
      ┌──────────────┐        ┌──────────────┐
      │Without written│        │ With written │
      │  Feedback     │        │  feedback    │
      └──────────────┘        └──────────────┘
             │                        │
             ▼                        ▼
        ⬭ Less                   ⬭ More
        incorporation            incorporation
```

Figure 1. Metacognitive training for peer feedback in L2 writing (* "metacognitive experience" appears in parenthesis as it has not been investigated in this study, although it has been argued to be a part of metacognition)

The findings of the present study suggest a few pedagogical implications for peer review in L2 writing, especially among young learners. First of all, the students appeared to develop an awareness of metacognitive knowledge and metacognitive regulations, which was consistent with the goals of the training provided to them, and which leads

to higher quality feedback and a smoother peer review interaction process. The students were able to provide a balanced range of content-related and language-related feedback instead of just surface language-related feedback. Secondly, metacognitive training in peer review assisted the learners in better understanding themselves, the task, as well as the relevant cognitive strategies, thereby leading to collaborative behaviors during the peer review interactions in class (Kong & Bui, 2019), and resulting in a willingness among the students to incorporate the peer feedback into their subsequent drafts. Thirdly, the unsatisfactory rate of revision resulting from the oral feedback, together with the high incorporation rate observed for the written comments, highlighted the importance of including further effective metacognitive strategies in the training for peer feedback (especially the ones associated with pre-feedback planning). It is, therefore, suggested that teachers should encourage their students to list the points related to their verbal interactions, which would later serve as reminders of the feedback and assist in their revisions.

The present study affirmed our hypothesis that it is possible to train learners (young and elders alike) to enhance their learning/peer review experience and outcomes by using metacognitive knowledge and metacognitive regulations (strategies), and thereby assisting them in becoming autonomous and self-regulated learners. However, it is noteworthy that these training programs should not be limited to peer review interactions. Teaching the students how to write down their comments effectively without disrupting the flow of interaction is also an important area for consideration. Future research could focus on investigating whether the feedback is of high quality and whether the incorporation of feedback leads to better writing quality. In addition, given the non-exploratory nature of the present study, experimental designs that allow direct testing of the metacognitive training in peer feedback may be applied in future investigations.

References

Abba, K. A., Zhang, S. S., & Joshi, R. (2018). Community college writers' metaknowledge of effective writing. *Journal of Writing Research*, *10*(1), 85-105. https://doi.org/10.17239/jowr-2018.10.01.04

Amores, M. J. (1997). A new perspective on peer-editing. *Foreign Language Annals, 30*(4), 513-522. https://doi.org/10.1111/j.1944-9720.1997.tb00858.x

Baker, L. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review, 1*(1), 3-38.

Bei, X. G. (2011). Formality in second language discourse: Measurement and performance. *Interdisciplinary Humanities, 28*(1), 32-41.

Bereiter, C., & Scardamalia, M. (2013). *The psychology of written composition*. New York, NY: Routledge.

Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing, 8*(3), 215-241. https://doi.org/10.1016/s1060-3743(99)80115-5

Bui, G. (2019). Task-readiness and L2 task performance across proficiency levels. In Z. Wen & M. Ahmadian (Eds.), *Researching L2 task performance and pedagogy: In honour of Peter Skehan*

(pp. 253-277). Amsterdam, the Netherlands: John Benjamins. https://doi.org/10.1075/tblt.13.12bui

Bui, G. (2018). A lexical approach to teaching formality in freshman L2 academic writing. In L. T. Wong & W. L. Wong (Eds.), *Teaching and learning English for academic purposes: Current research and practices* (pp.111-124). New York, NY: Nova Science Publishers, Inc.

Bui, G., & Huang, Z. (2018). L2 fluency as influenced by content familiarity and planning: Performance, methodology and pedagogy. *Language Teaching Research*, *22*(1), 94-114. https://doi.org/10.1177/1362168816656650

Bui, G., & Yu, R. (2019). Spaced multi-draft composing and feedback in mainland Chinese English as a foreign language secondary school writing literacy. In B. Reynolds & M. F. Teng (Eds.), *English literacy instruction for Chinese speakers* (pp. 127-141). Singapore: Palgrave Macmillan. https://doi.org/10.1007/978-981-13-6653-6_8

Caulk, N. (1994). Comparing teacher and student responses to written work. *TESOL Quarterly, 28*(1), 181-188. https://doi.org/10.2307/3587209

Chamot, A. U. (2005). Language learning strategy instruction: Current issues and research. *Annual Review of Applied Linguistics, 25*, 112–130. https://doi.org/10.1017/s0267190505000061

Chang, C. Y. H. (2015). Teacher modeling on EFL reviewers' audience-aware feedback and affectivity in L2 peer review. *Assessing Writing, 25*, 2-21. https://doi.org/10.1016/j.asw.2015.04.001

Chang, Y. H. (2016). Two decades of research in L2 peer review. *Journal of Writing Research, 8*(1), 81-117. https://doi.org/10.17239/jowr-2016.08.01.03

De Guerrero, Maria, C. M., & Villamil, O. S. (1994). Social-cognitive dimensions of interaction in L2 peer revision. *The Modern Language Journal, 78*(4), 484-496. https://doi.org/10.1111/j.1540-4781.1994.tb02065.x

Dixon, H., & Hawe, E. (2017). Creating the climate and space for peer review within the writing classroom. *Journal of Response to Writing, 3*(1), 6–30.

Faigley, L., & White, S. (1981). Analyzing revision. *College Composition and Communication*, *32*(4), 400-414.

Ferris, D., & Hedgcock, J. (2014). *Teaching L2 composition: Purpose, process, and practice* (3rd ed.). New York, NY: Routledge.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–236). Hillsdale, NJ: Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34(10)*, 906 - 911. https://doi.org/10.1037/0003-066x.34.10.906

Hansen, J. H., & Liu, J. (2005). Guiding principles for effective peer response. *ELT Journal, 59*(1), 31-38. https://doi.org/10.1093/elt/cci004

Hirose, K. (2012). Comparing written-only and written-plus-spoken peer feedback in a Japanese EFL university context. *Asian Journal of English Language Teaching, 22*, 1-23.

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing, 7*(3), 255-286. https://doi.org/10.1016/s1060-3743(98)90017-0

Kong, A. (2019). A dual case study of perceptions of strategy-based training and peer review stances among Hong Kong secondary students: An activity theory perspective. In B. L. Reynolds & F. Teng (Eds), *Innovative approaches in teaching writing to Chinese speakers*. Berlin: De Gruyter.

Kong, A., & Bui, G. (2019). Reader stances and writer responses in L2 peer review: A study of L2 writing literacy among Hong Kong secondary school students. *Asian EFL Journal, 21*(3), 140-186.

Lee, I. (2017). *Classroom assessment and feedback in L2 school contexts*. Singapore: Springer. https://doi.org/10.1007/978-981-10-3924-9

Lee, I., & Mak, P. (2018). Metacognition and metacognitive instruction in L2 writing. *TESOL Quarterly*, *52*(4), 1085-1097. https://doi.org/10.1002/tesq.436

Liu, J., & Hansen, J. G. (2002). *Peer response in second language writing classrooms*. Ann Arbor: University of Michigan Press.

Liu, E. Z. F., & Lin, S. S. J. (2007). Relationship between peer feedback, cognitive and meta-cognitive strategies and achievement in networked peer assessment. *British Journal of Educational Technology, 38*(6), 1122-1125. https://doi.org/10.1111/j.1467-8535.2007. 00702.x

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

Min, H. T. (2006). Training students to become successful peer reviewers. *System, 33*(2), 293 – 308. https://doi.org/10.1016/j.system.2004.11.003

Nelson, G. L., & Carson, J. G. (1998). ESL students' perceptions of effectiveness in peer response groups. *Journal of Second Language Writing, 7*(2), 113-131.https://doi.org/10.1016/s1060-3743(98)90010-8

Patchan, M. M., & Schunn, C. D. (2016). Understanding the effects of receiving peer feedback for text revision: Relations between author and reviewer ability. *Journal of Writing Research, 8*(2). https://doi.org/10.17239/jowr-2016.08.02.03

Paulus, T. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing, 8*(3), 265-289. https://doi.org/10.1016/s1060-3743(99)80117-9

Qin, L., & Zhang, L.J. (2019). English as a foreign language writers' metacognitive strategy knowledge of writing and their writing performance in multimedia environments. *Journal of Writing Research, 12*(2), 393-413. doi:10.17239/jowr-2019.11.02.06

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. https://doi.org/10.1177/1362168812459151

Ren, H., & Hu, G. (2012). Peer review and Chinese EFL/ESL student writers. *English Australia Journal, 27*(2), 3.

Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT journal, 59*(1), 23-30. https://doi.org/10.1093/elt/cci003

Ruegg, R. (2015). The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Linguistics and Education, 29*, 73-82. https://doi.org/10.1016/j.linged.2014.12.001

Saliu-Abdulahi, D., Hellekjær, G. O., & Hertzberg, F. (2017). Teachers' (formative) feedback practices in EFL writing classes in Norway. *Journal of Response to Writing, 3*(1), 31–55.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional science, 26*(1-2), 113-125.

Scott, B. M., & Levy, M. G. (2013). Metacognition: Examining the components of a fuzzy concept. *Educational Research, 2*(2), 120–131. https://doi.org/10.5838/erej.2013.22.04

Sengupta, S. (1998). Peer Evaluation: 'I am not the teacher'. *ELT Journal, 52*(1), 19-28. https://doi.org/10.1093/eltj/52.1.19

Stanley, J. (1992). Coaching student writers to be effective peer evaluators. *Second Language Writing, 1*(3), 217-233. https://doi.org/10.1016/1060-3743(92)90004-9

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: two adolescent French immersion students working together. *The Modern Language Journal, 82*(3), 320-337. https://doi.org/10.1111/j.1540-4781.1998.tb01209.x

Teng, F. (2016). Immediate and delayed effects of embedded metacognitive instruction on Chinese EFL students' English writing and regulation of cognition. *Thinking Skills & Creativity, 22*, 289-302. https://doi.org/10.1016/j.tsc.2016.06.005

Teng, F. (2017). The effects of task-induced involvement load on word learning and confidence judgments mediated by knowledge and regulation of cognition. *Educational Sciences: Theory & Practice, 17*(3), 791-808. https://doi.org/10.12738/estp.2017.3.0167

Teng, F. (2019). Tertiary-level students' English writing performance and metacognitive awareness: A group metacognitive support perspective. *Scandinavian Journal of Educational Research.*

https://doi.org/10.1080/00313831.2019.1595712

Teng, F., & Huang, J. (2019). Predictive effects of writing strategies for self-regulated learning on secondary school learners' EFL writing proficiency. *TESOL Quarterly, 53*(1), 232-247. https://doi.org/10.1002/tesq.462

Tsui, A., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing, 9*(2), 147 – 170. https://doi.org/10.1016/s1060-3743(00)00022-9

Van Velzen, J. H. (2016). *Metacognitive learning: Advancing learning by developing general knowledge of the learning process*. Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-24433-4_4

Villamil, O. S., & De Guerrero, M. C. (1996). Peer revision in the L2 classroom: Social-cognitive activities, mediating strategies, and aspects of social behavior. *Journal of Second Language Writing, 5*(1), 51-75. https://doi.org/10.1016/s1060-3743(96)90015-6

Zhang, S. (1995). Reexamining the affective advantage of peer feedback in the ESL writing class. *Journal of second language writing, 4*(3), 209-222. https://doi.org/10.1016/1060-3743(95)90010-1

Zhang, L. J. (2001). Awareness in reading: EFL students' metacognitive knowledge of reading strategies in an acquisition-poor environment. *Language Awareness, 10*(4), 268-288. https://doi.org/10.1080/09658410108667039

Zhang, L. J. (2010). A dynamic metacognitive systems account of Chinese university students' knowledge about EFL reading. *TESOL Quarterly, 44*(2), 320–353. https://doi.org/10.5054/tq.2010.223352

Zhu, W. (1995). Effects of training for peer response on students' comments and interaction. *Written Communication, 1*(4), 492-528. https://doi.org/10.1177/0741088395012004004

Appendix A: Sample Interview and Stimulated Recall Questions

Guidelines for the Pre-interview

1. Do you have any earlier peer review experience? If yes, can you please briefly describe your experience?
2. What do you think of peer review? Do you think it is useful?

Guidelines for the Post-interview

1. What are your perceptions of peer review? Do you think it is useful? Why?
2. What strategies did you prefer to use during peer review? Why?
3. Do you think it is important to give both the global comments and surface comments to your peers? And what kinds of comments did you prefer to give/receive? Why?
4. Did you write down your comments? Why?
5. Did you revise your writing drafts according to the peer comments? Why?
6. Do you think our training is useful in facilitating the peer review process?

Guidelines for Stimulating Recalls (with reference to the individual peer review recordings)

1. Why did you behave in certain ways (e.g., intensive/loose interactions, the use of L1, etc.) in the last peer review session?
2. Did you have any difficulties interacting with the peer?