

Lower English proficiency means poorer feedback performance?: A mixed-methods study

Abstract: This study adopts a mixed-methods design and examines the relation between English proficiency and peer feedback performance. Data sources included peer feedback made by 23 lower English proficiency (LEP) students and 23 higher English proficiency (HEP) students, and semi-structured interviews with four LEP and four HEP students from that sample. Quantitative analysis did not find significant difference between the two groups in feedback amount or feedback quality, but significant difference was found in feedback type. LEP students tended to make more clarification requests in content feedback, more suggestions and fewer direct corrections in language feedback. Qualitative analysis of students' interviews showed that individual factors (i.e. genre knowledge, L1, L2, reference materials, and imagined identities) and contextual factors (i.e. anonymity, feedback separation, and time availability) shaped their feedback processes. Based on these findings, a conceptual framework is suggested to explain how feedback performance is enabled or constrained by cognitive, affective, sociocultural, and instructional factors. The framework can be a useful heuristic for EFL teachers to create facilitative conditions to engage and empower LEP students in feedback activities.

Keywords: English proficiency; peer feedback; mixed-methods design; feedback performance

1. Introduction

Peer feedback is well supported by a number of theories, including process writing theory, interactionist theory, sociocultural theory of learning, and many others (Hyland & Hyland, 2006; Lee, 2017; Yu & Lee, 2016a). Empirically, peer feedback has been found to reduce language errors (Diab, 2010), improve text complexity (Allen & Mills, 2013), increase audience awareness (Chang, 2015), and develop learner autonomy (Yang, Badger, & Yu, 2006). Studies further show that peer feedback is not inferior to teacher feedback (Caulk, 1994), but differs in focal areas. Specifically, peer feedback tends to focus more on content and meaning issues, while teacher feedback focuses on language issues (Tsui & Ng, 2010;

Yang et al., 2006). Therefore, peer feedback and teacher feedback can play complementary roles in EFL writing pedagogies (Lee, 2017; Tsui & Ng, 2000).

Despite the theoretical and empirical support of peer feedback, EFL writing instructors are faced with multiple challenges when designing feedback activities. One vexing challenge is students' English proficiency (Nelson & Murphy, 1993; Tsui & Ng, 2000). When we assign students into feedback groups, should we form heterogeneous or homogenous groups? While ability matching is not a significant predictor of feedback quality in the L1 context (Huisman et al., 2017), EFL students hold mixed perceptions of their ability to give feedback. For instance, in Allen's (2015) survey study, half of the participants believed that English proficiency had strong or some influence on the feedback process, while the other half did not think so. In Zhang's (2011) study, students regarded English proficiency as a prominent factor impacting their feedback focus, type, and accuracy. Due to their under-developed English competence, lower English proficiency (LEP) students are usually considered not knowledgeable enough to identify or rectify language issues (Nelson & Murphy, 1993). Consequently, they tend to be marginalized and play the role of a feedback receiver (Lee, 2017; Yu & Lee, 2016a). Is this distrust of LEP students justified? Unfortunately, not much is known given the paucity of empirical research on LEP students' feedback products and processes (Yu & Lee, 2016b).

Existing studies examining the relation between English proficiency and feedback performance can be grouped into two strands. The first strand compares how higher English proficiency (HEP) students and LEP students differ in their feedback products. Studies have tended to employ a quantitative design and yield conflicting results. For instance, Kamimura (2006) found that LEP students made more comments than HEP students, while Allen and Mills (2016) found that HEP students gave more comments than LEP students. The result discrepancy might be caused by the grouping methods and feedback procedures. In Kamimura's (2006) study, LEP and HEP students swapped the drafts amongst themselves and formed matched proficiency dyads (LEP-LEP; HEP-HEP). They were asked to respond to a feedback worksheet, which contained more content-related questions than language-related questions. As such, LEP students' attention might be more drawn to their LEP peers' content issues (which were more prevalent than those in HEP peers' drafts), thus boosting the total number of peer comments.

The second line of research adopts a qualitative design to describe the dynamics in mixed-proficiency feedback groups. Yu and Lee (2016b) focused on 12 students in three heterogeneous groups and found that LEP students were able to provide a wide range of quality comments, about 80% of which were subsequently incorporated by their peers. The researchers identified four factors that empowered LEP students in peer feedback: congenial group relationship, positive attitude, use of first language (L1), and feedback training. The benefits of heterogeneous grouping is corroborated by Yu and Hu (2017a). HEP students could learn from the feedback given by LEP students, and developed their audience awareness and L2 writing knowledge. Similarly, Allen and Katayama (2016) examined the group dynamics of six dyads with different English proficiency levels. After triangulating multiple data sources, the researchers concluded that language proficiency and *perception* of relative proficiency had a great impact on the number and the type of feedback provided and incorporated.

Although the aforementioned studies are informative, their insights are potentially limited by the research methods. Quantitative methods have been used to measure the *products* of peer feedback, such as the number, type, quality, and uptake of feedback made by LEP students (Allen & Mills, 2016; Kamimura, 2006). On the other hand, qualitative methods have been adopted to understand the *processes*, whereby LEP students negotiate a multitude of individual and contextual factors (Allen & Katayama, 2016; Yu & Lee, 2016; Yu & Hu, 2017a, b). However, feedback product and process are two sides of the same coin—feedback performance. To fully understand how LEP students perform and why they perform the way they do, the connection between product and process needs to be addressed. Therefore, a mixed-methods design is warranted to examine whether quantitative differences exist in the feedback products of HEP and LEP students, *and* what factors shape these differences or similarities in the processes of giving feedback.

Another potential limitation of previous research is that LEP and HEP students did not work on the *same* piece of writing. In Kamimura's (2006) study, as described previously, matched proficiency dyads were created, while in Allen and Mills' (2016) study, students formed self-initiated dyads, producing four types of writer-reviewer combinations: HEP-HEP, HEP-LEP, LEP-HEP, and LEP-LEP. In all these scenarios, students worked on different essays, which had varying amounts of treatable issues for feedback-givers. A

more refined design is to have a LEP student and a HEP student make feedback on the same piece of writing by a third student. In this way, we place LEP and HEP students on a level playing field, and we are more justified to associate the observed different feedback performances with the proficiency gap.

2. The study

2.1 Research questions

To gain a fuller understanding of the relation between English proficiency and peer feedback performance, this study was guided by the following questions: (1) Do HEP and LEP students differ in feedback amount, type, and quality? (2) Why do HEP and LEP students share similarities or differences in feedback performance?

2.2 Context and participants

This study involved 69 first-year students who studied English as a foreign language (EFL) at a first-tier university in Guangdong Province, southern China. They came from three parallel classes and were enrolled in a *Narrative Writing* course focusing on such narrative skills as plot, characterization, and point of view. At the beginning of the semester, they took two sessions of computer-mediated feedback training, in which they modeled from the instructor and used a word processor (Microsoft Word) to make feedback on a sample story. The training also elucidated (a) concepts of content feedback and language feedback, (b) roles, functions and strategies of peer feedback, and (c) basic word processing skills. After the two training sessions, students commented on another sample story at home and met with the instructor the next week in a 20-minute teacher-student conference. This step was to ensure that students were clear about the goal and procedure of giving feedback (Min, 2005; Rahimi, 2013). Throughout the course, students wrote three short stories as assignments, all of which were peer reviewed. This study focused on the final project, which required students to write a story of about 800 words. There was no specific prompt except that students needed to demonstrate the knowledge and skills of narrative writing covered in the course. The project was worthy of 30% of the course grade, with the first draft taking up 5%, peer feedback 5%, and the revised draft 20%. For the purposes of this

study, peer feedback was anonymously conducted in a computer lab (see the next subsection for detail).

2.3 Procedure

The study followed a mixed-methods design and was conducted in four stages. In the first stage, 69 students submitted their first drafts, and on the same day took Oxford Quick Placement Test (Oxford University, 2001) in class. The students were then rank-ordered on the test score as an indicator of their English proficiency. Next, a colleague with 10 years of teaching narrative writing and I independently rated the first drafts, using a five-level rubric (from “poor” to “excellent”) with five criteria (i.e. narration, description, coherence, language accuracy, and language style). The inter-rater reliability was measured by Spearman correlation and found to be very high ($r = 0.91$, $p < 0.001$). The average of the two raters’ scores was taken as the grade of the first drafts. To provide a common ground to compare HEP and LEP students’ feedback performance, the 23 middle-ranked stories were selected as materials of this study. A “moving window” procedure was adapted from Huisman et al. (2017) in that each middle-ranked story was assigned to a pair of students—one from the top 33% (i.e. the HEP group) and the other from the bottom 33% (i.e. the LEP group), thus “keeping the [proficiency] difference...as constant as possible” (p. 1438). The remaining 46 stories were randomly assigned to the remaining students, which were not the focus of this study. After that, students’ names were all replaced by four-digit numbers to ensure anonymity (Lundstrom & Baker, 2009). The previous three steps were facilitated by the functions in Microsoft Excel 2016 (e.g. MATCH, INDEX, and RANDBETWEEN).

In the second stage (one week later), peer review was conducted in a computer lab outside the regular class time. The session had a designated start time but not a definite end time, so the students could take all the time they needed to provide feedback. Students were instructed to provide content feedback first and then language feedback in the forms of Track Changes and marginal comments in Microsoft Word. They could use L1, L2 or a combination of both to give feedback. Blackboard, the course management system, was used to facilitate the distribution and collection of drafts (see Appendix A for a step-by-step worksheet).

In the third stage, immediately after the lab session, students received the peer-reviewed e-copies and were given one week to revise the draft. Initially, I intended to examine feedback uptake by comparing the first and revised drafts. I used the “Compare” function in Microsoft Word 2016 to display the changes between the drafts. I then manually coded the changes in 30% of the revised drafts and found that students substantially reworked their stories, such as altering the sequence of events and removing protagonist description. It was not possible to trace whether the authors had accepted or rejected the suggestions made by their peers. Therefore, in the coding process, I focused on the feedback amount, quality, and type, but not uptake. Coding protocols will be reported in Section 2.4.

In the fourth stage, to allow free expression of their thoughts, students were not approached for an interview until the course grades were finalized in the course management system. Eight students were approached as potential interviewees based on the principle of maximum variation (Patton, 2002), such that the top 1, 25th percentile, 50th percentile, 75th percentile from the HEP group and their counterparts from the LEP group were selected. They all consented to be interviewed and their information is reported in Table 1. The interview included stimulated recall, in which each participant was given the draft he or she peer-reviewed and asked to recall the feedback process. I also went over a prepared list of questions generated from the feedback pattern I observed in the student’s reviewed draft (see Appendix B for an example of interview outline). The interview was conducted in Cantonese, the first language of the participants. The interview extracts cited in this paper were translated by the researcher and checked by the aforementioned colleague.

Table 1. Profile of eight interviewees.

Pseudonym	Gender	English Proficiency (Oxford Quick Placement Test score)
Wong	Female	Higher (36)
Yip	Female	Higher (36)
Lee	Male	Higher (35)
Tam	Female	Higher (33)
Chan	Female	Lower (28)
Tsai	Female	Lower (27)

Chow	Female	Lower (27)
Kwok	Male	Lower (26)

2.4 Data analysis

The quantitative analysis examined whether the independent variable (English proficiency) had an impact on the dependent variables (feedback amount, type, and quality), thereby addressing the first research question (RQ-1). Table 2 summarizes the variables and their measurements, and the following paragraphs describe how each variable was measured in greater detail.

Table 2. Summary of independent and dependent variables.

Variable	Measurement
Independent variable	
English proficiency	Oxford Quick Placement Test (Oxford University, 2001)
Dependent variables	
Feedback amount	The total number of marginal comments and direct changes made in the Word files
Feedback type	A four-type protocol for content feedback and a six-type protocol for language feedback
Feedback quality	Min's (2006) method of assessing feedback quality: original better, suggestion better, and no difference

English proficiency was assessed by Oxford Quick Placement Test (Oxford University, 2001). In this study, only the first part of the test was administered and the maximum score was 40. The top 23 students and the bottom 23 students were regarded as the HEP group ($M = 34.52$, $SD = 1.16$) and the LEP group ($M = 26.74$, $SD = 0.75$), which were equivalent to the CEFR B2 and B1 levels. Independent-samples t test confirmed that the two groups statistically differed in English proficiency ($t(44) = 26.96$, $p < .001$).

Feedback amount was measured by the summated number of marginal comments and direct changes made in a Word file. The number of marginal comments was taken from

the Microsoft Word report, while the number of direct changes was manually counted. It should be noted that regardless of the size of the change (Min, 2006) it was counted as one change. For instance, in the following example (see Figure 1), ten changes were counted.

When I ~~stand~~¹stood in the bushes, I saw Lola and John standing by the ~~artificial~~² lake on the campus. The surrounding was so silent ~~and~~³that I ~~could~~⁴ heard my heart ~~was~~⁵ thumping. I saw Lola passing on something like a letter to John. That must be a love letter! I clenched my fists and murmured, "Don't ~~receiv~~⁶take it. Don't ~~receive~~⁷!" But I saw John grinning and taking the letter. I howled in my heart, "How could John take Lola's love? Though we ~~didn't have a chat~~⁸haven't talked to each other, ~~but~~⁹ I can feel that he ~~falls~~¹⁰is in love with me. Every day he would smile to me so brightly. Maybe he also mistakes me and Lola. We are twin sisters. "

Figure 1. An example of counting feedback amount in a student's reviewed draft.

Feedback type was coded by the same colleague and me. We first independently coded 30% of the peer-reviewed drafts. We then discussed the codes and referred to feedback types reported in the literature (Allen & Mills, 2016; Min, 2005). After deliberation, we agreed to code content feedback in four categories (i.e. Suggestion, Clarification Request, Problem Description, and Praise) and code language feedback in six categories (i.e. Suggestion, Clarification Request, Problem Description, Metalinguistic Explanation, Direct Correction, and Praise). Using these protocols, we independently coded all the feedback. The percentage for inter-rater agreement was 97% for content feedback and 94% for language feedback. We discussed the inconsistent codes until consensus was reached.

Feedback quality was measured using Min's (2006) method: original better, suggestion better, and no difference. The same colleague and I independently rated the students' marginal comments and direct changes. The percentage agreement was 93% for content feedback, and 89% for language feedback. The inconsistencies were resolved by discussion.

As some data were not normally distributed, I used the Wilcoxon signed-rank test, the non-parametric equivalence of paired-samples *t* test (Larson-Hall, 2010), to compare the feedback amount, type, and quality between HEP and LEP students. All the statistical analyses were conducted in SPSS 21.

To answer the second research question (RQ-2), I followed a grounded theory approach to analyze the interview transcripts. In open coding (Glaser, 1992), I first marked up all the

sections where the students explained the factors that shaped their feedback behaviors. I then grouped the marked-up sections based on themes (i.e. feedback amount, type, and quality). In theoretical coding (Glaser, 1992), I compared and merged explanations in the same theme and generalized factors/reasons to form coherent categories. Then, I triangulated the factors/reasons with the quantitative data and made further adjustments. Finally, I was able to identify eight reasons, which were classified into four factors and two domains (see Section 4 for detail).

3. Research results

3.1 RQ-1: Do HEP and LEP students differ in feedback amount, type, and quality?

3.1.1 Feedback amount

As shown in Table 3, the two proficiency groups did not statistically differ in the total amount of content or language feedback. This result is different from the observation in Allen and Mills (2016), who found that HEP students made more comments. The inconsistency might be caused by the grouping method and the policy of anonymity. In this study, an HEP student and a LEP student commented on the same text and a double-blind policy was enacted. In Allen and Mills’s (2016) study, students formed self-initiated, identifiable dyads. As such, it was possible that LEP students made fewer comments when paired with HEP students and HEP students made more comments when paired with LEP students because of the *perceived* proficiency gap. The result in this study is also inconsistent with Kamimura’s (2006) observation that LEP students made more comments than HEP students. In Kamimura’s study, only matched proficiency dyads were formed, and the students’ attention was drawn to the content issues, which were more prevalent in LEP students’ essays, thereby increasing the total amount of feedback made by LEP students.

Table 3. Feedback amount: descriptive statistics and Wilcoxon signed-rank tests.

Feedback amount	HEP M (SD)	LEP M (SD)	Z	P
Content feedback	9.48 (4.47)	8.22 (3.41)	-1.12	0.27
Language feedback	23.52 (12.65)	20.22 (11.84)	-1.01	0.29

3.1.2 Feedback type

As Table 4 shows, HEP and LEP students differed in only three types of feedback. Specifically, LEP students made more clarification requests when giving content feedback ($Z = -2.89$, $p < 0.01$). They also made more suggestions ($Z = -2.16$, $p = 0.03$) and fewer direct changes ($Z = -2.02$, $p = 0.04$) in language feedback. Other than that, the two groups did not display statistical differences in feedback type.

Table 4. Feedback type: descriptive statistics and Wilcoxon signed-rank tests.

Feedback type	HEP M (SD)	LEP M (SD)	Z	P
Content				
Suggestion	2.61 (2.69)	1.61 (1.56)	-1.57	0.12
Clarification request	0.65 (0.98)	2.13 (1.87)	-2.89	0.004**
Problem description	3.35 (2.72)	2.22 (2.01)	-1.21	0.23
Praise	2.87 (2.65)	2.65 (3.10)	-0.59	0.56
Language				
Suggestion	1.78 (2.17)	3.91(4.67)	-2.16	0.03*
Clarification request	5.96 (3.39)	4.39 (2.93)	-1.67	0.10
Problem description	1.74 (2.16)	2.04 (2.34)	-0.44	0.66
Metalinguistic	0.65 (1.11)	0.78 (1.41)	-0.28	0.78
Explanation				
Direct correction	10.87 (9.33)	7.13 (6.76)	-2.02	0.04*
Praise	2.52 (3.37)	1.96 (2.51)	-0.45	0.66

* $p < 0.05$; ** $p < 0.01$

3.1.3 Feedback quality

As Table 5 shows, the two groups display no statistical difference in any of the six variables. This indicates that HEP and LEP students did not differ in feedback quality, suggesting

that HEP and LEP students were able to assume the dual roles of contributors and beneficiaries in feedback activities (Yu & Hu, 2017a; Yu & Lee, 2016b).

Table 5. Feedback Quality: descriptive statistics and Wilcoxon signed-rank tests.

Feedback quality	HEP M (SD)	LEP M (SD)	Z	P
Content feedback				
Suggestion better	7.78 (3.38)	6.65 (2.72)	-1.68	0.09
No difference	1.43 (1.88)	1.30 (2.22)	-0.68	0.49
Original better	0.22 (0.52)	0.26 (0.68)	-0.32	0.75
Language feedback				
Suggestion better	20.61 (11.8)	17.30 (12.23)	-1.25	0.21
No difference	2.26 (1.18)	2.13 (2.05)	-0.22	0.82
Original better	0.65 (0.83)	0.78 (1.24)	-0.99	0.92

3.2 RQ-2: Why do HEP and LEP students share similarities or differences in feedback performance?

3.2.1 Feedback amount

The quantitative analysis shows that the two cohorts did not differ in feedback amount. In the interviews, when asked why they were able to make a substantial amount of feedback, LEP students alluded to three reasons: anonymity, L1, and imagined identity.

First, LEP students pointed out that they felt liberated to express themselves because of the double-blind policy. As Chow commented, “keeping us anonymous is wonderful. I can give my comments without worrying about hurting the feelings of my peers.” Kwok further elaborated that “because of the face issue, we tend not to be frank with each other. This time, I feel like I am given a free rein. I can be completely honest in the review.” As such, anonymity encouraged critical and deeper engagement with the texts, thereby contributing to feedback performance.

Additionally, LEP students cited L1 as an empowering factor. Chan explained that “I am so glad that we can use Chinese to make comments. I can express my views in a more precise way and make myself understood.” Similarly, Tsai noted that L1 allowed her to be

articulate in feedback: “When I comment in Chinese, I am able to give more comments, especially in content feedback. [...] If I had to give feedback in English, I would be tongue-tied.” Clearly, L1 became a powerful artifact that enabled LEP students to fully express themselves and make detailed, incisive feedback.

Third, LEP students made great effort to provide feedback because they wanted to act as a helpful and constructive reviewer. As Chan pointed out, “I don’t want to project a sloppy image. So, I follow a self-imposed rule: on every page, I must come up with at least one comment.” Similarly, Chow’s imagined identity motivated her to give a large amount of feedback: “in this semester, I’ve experienced three peer reviewers, and one of them was unhelpful. I don’t want to become someone I dislike, so I constantly remind myself to give as many helpful comments as possible.”

Contrary to our expectations, HEP students did not make more feedback than LEP students. In the interviews, HEP students opined that they did not point out self-evident issues because their peers should have the ability to recognize them. For instance, when asked why he did not treat the mistakes highlighted in red and blue lines by Microsoft Word, Lee responded that “I don’t want to correct those obvious mistakes. As they are already marked up by the software, I trust that my peers are able to correct them on their own.” Tam further remarked that “if we correct self-evident issues, we are treating the authors as primary school students. We are adults now. We can think for ourselves.” It appeared that HEP students had confidence in their peers’ abilities to rectify obvious issues, so they did not treat them in peer feedback.

3.2.2 Feedback type

As the quantitative results show, LEP students made more clarification requests in content feedback, more suggestions and fewer direct changes in language feedback. Taken together, LEP students appeared to be less assertive than HEP students. In the interviews, LEP students confirmed that they “are more suggestive than aggressive when giving feedback” (Chan). They further explained that they asked the author to clarify his/her intentions, because their previous experience reminded them to keep an open mind:

In the previous assignment, I experienced one reviewer. Totally impatient! He did not make any effort to understand my plotline. In this project, I asked clarification

questions as an attempt to understand my peer's story, instead of jumping into conclusion that it was bad. (Chow)

When asked why they made more suggestions but fewer direct changes in language feedback, they cited their English proficiency as a main reason. As Kwok confessed, "my English is not very good, so I don't want to be adamant. Unless I am absolutely sure, I don't make direct language changes. Rather, I use marginal comments to suggest alternative versions." Two LEP students further remarked that making marginal comments instead of direct changes would seem "less intrusive" (Tsai) and "non-threatening" (Chan).

Conversely, HEP students made more direct language changes because their English proficiency afforded them the confidence to do so. For instance, Lee stated that "There is nothing to suggest. It is wrong, so I correct it." Additionally, HEP students explained that they made direct language changes where they thought their peers were unable to do so: "I know tense is a thorny issue in story writing. My peers usually confuse tenses. So I made direct corrections for them" (Wong).

When asked why they made fewer clarification requests in content feedback, HEP students said that their objective was to externalize their feelings as a reader. For instance, Yip focused more on her reaction to the writing than what the story tried to say: "My job as a reviewer is to read and react. I note down my personal feelings when I read my peers' stories: exciting, funny, abrupt, choppy, and so on. The author can use these responses to improve the story." Tam held a more radical opinion and placed little importance on checking writer's intention: "Isn't reader response about RESPONSE? What does author's intention have to do with it? Roland Barthes proclaimed the death of the author. Reader's impression is more important than author's intention!" Driven by the keen desire to share the emotive responses, HEP students tended not to ask peer writers to clarify their intended meanings.

3.2.3 Feedback quality

The quantitative analysis showed that HEP and LEP students did not differ in either content or language feedback quality. In the interviews, when asked why they were able to make effective content feedback, LEP students explained that they were familiar with the genre. As Chan explained, "I am quite interested in reading short stories. I have read hundreds of

them before this course. I have no problem detecting plot issues.” LEP students’ ability to provide content feedback was also recognized by HEP students. For instance, Yip believed that “even though some students are less proficient in English, they know what makes a great story.” Wong also noted that “we’ve received systematic training about narrative writing in this semester, so English proficiency does not matter much in content feedback.”

As regards language feedback, LEP students resorted to online reference materials to enhance their feedback quality. For instance, Kwok made sophisticated use of online dictionaries. He used different dictionaries for decoding meanings in content feedback and encoding meanings in language feedback:

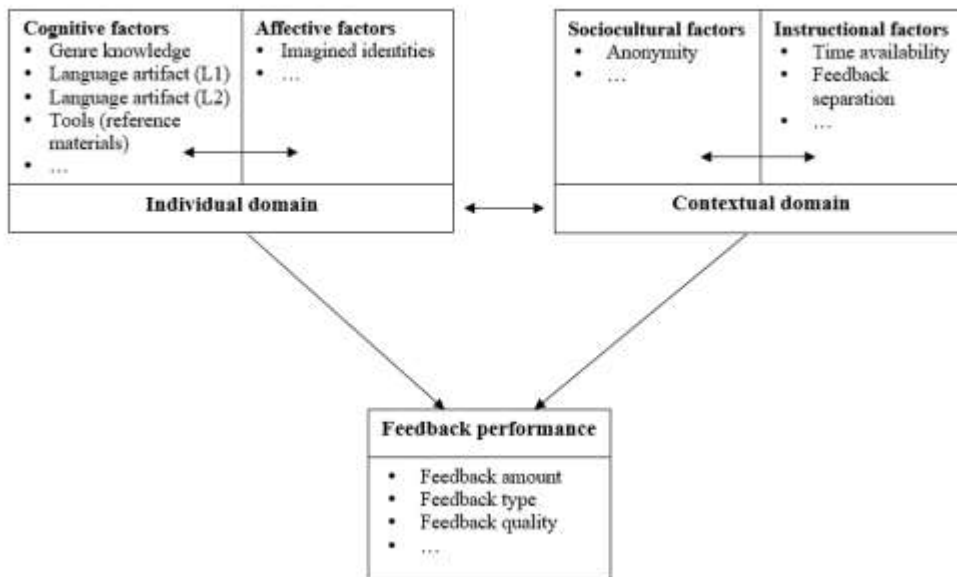
I use two digital dictionaries. With the first one, I just need to place the cursor on the word and the meanings will instantly pop up. I use this function when giving content feedback. When giving language feedback, I use *Oxford Advanced Learners’ Dictionary*. I check collocations and other grammatical information.

Similarly, Tsai was a mindful user of Corpus of Contemporary American English (COCA), reporting that “When I read some strange or unfamiliar expressions, I look them up in COCA to check whether they are idiomatic. If not, I would point them out.” These extracts show that LEP students were able to garner technology artifacts to compensate their English proficiency in peer feedback.

Additionally, LEP students cited two instructional reasons that contributed to their feedback quality. First, content feedback was separated from language feedback, which set a particular focus for them. Chan remarked that “I directed my attention to content or language at a given time. This helped me better detect potential problems.” Tsai also noted that “if we mixed content and language feedback, we had to switch back and forth between the two. I was not good at multi-tasking.” Second, no time pressure was imposed in the feedback session, so students could take time to access information and think carefully before giving feedback. Chow recalled that “it took me over 1.5 hours to comment on one story, because I consulted online dictionaries dozens of times to check grammatical information.” Kwok concurred that “I felt relaxed because I could take all the time I needed to think really hard about the story and my comments.” These extracts show that feedback separation and time availability were two instructional affordances that enabled LEP students to enhance their feedback quality.

4. Discussion

Based on the quantitative and qualitative analysis, a conceptual framework is suggested to understand the feedback performance by LEP and HEP students (see Figure 2). Inspired by Ellis (2010) and Yu and Hu (2017b), the framework has two overarching domains: individual and contextual. Individual factors include cognition and affection, while contextual factors include sociocultural and instructional affordances and constraints. Each factor has its specific instantiations and impacts on feedback performance. The following subsections first explain each of the four factors and then discuss the interplay of factors and domains.



Notes: Double ended arrows indicate that feedback performance is usually co-determined by an interplay of factors within and across the individual and contextual domains. Ellipses indicate that more items are possible in addition to the factors and variables observed in this study.

Figure 2. A conceptual framework of factors influencing feedback performance.

4.1 Cognitive factors

The students reported four reasons as important cognitive factors that influenced their feedback performance: genre knowledge, L1, L2, and reference materials.

Familiarity with genre features can reduce the cognitive demand when processing genre texts. Previous research has shown that students are less confident in giving peer feedback

if they deal with an unfamiliar genre. For example, students in Zhao's (2011) study confessed that their ability to give feedback was limited by the genre knowledge of English poems. In this study, however, students had sufficient prior exposure to narrative writing and after one-semester training they developed systematic knowledge of short stories. That was why they could provide useful content feedback when they commented on their peers' stories. Taken together, it can be followed that systematic genre knowledge can facilitate LEP students to provide quality content feedback.

A language is an important mediational artifact in the process of sociocultural learning (Vygotsky, 1978). Traditionally, the use of L1 has been discouraged (if not forbidden) in EFL classrooms. This English-only policy, to some extent, linguistically disables LEP students in peer feedback. In this study, the use of L1 was appreciated by LEP students as contributing to their feedback amount. This adds further empirical evidence to the observation that the use of L1 encourages LEP students' engagement in feedback activities (Yu & Hu, 2017a, b; Yu & Lee, 2016b). L1, as an empowering language artifact, allows students to perform higher-order cognitive tasks, such as analyzing problems, evaluating different options, and articulating solutions. It gives a voice to LEP students, who might otherwise be voiceless in L2.

Following from the previous point, L2, also a language artifact, has its impact on the amount and type of feedback. As HEP students reported, their English proficiency led them to believe that correcting obvious language mistakes (such as those already highlighted by Microsoft Word) was unnecessary because the author should have the ability to treat them. Thus, they refrained from making comments on the obvious issues. Additionally, L2 proficiency mediated the way both groups provided feedback. For LEP students, their lack of confidence in English proficiency led them to make more language suggestions in marginal comments and fewer direct corrections. Conversely, HEP students, more confident in their English proficiency, made a larger amount of direct changes where they thought their peers were unable to treat grammar issues.

Material tools are another mediational means in the process of sociocultural learning (Vygotsky, 1978). As described by LEP students, they made frequent and sophisticated use of online reference materials (e.g. COCA and *Oxford Advanced Learners' Dictionary*) to check meanings, collocations, and grammatical information. This indicates that in

computer-mediated peer feedback the relative disadvantage of English proficiency can be compensated by acting upon the affordances of technology and information artifacts. In Yoon's (2016) words, these reference materials were "effective cognitive [tools] extending the cognitive powers of the participants in solving lexical and grammatical problems" (p. 209). If the students had otherwise conducted peer feedback in a traditional classroom with no access to computer or Internet, they would be unable to locate and utilize artifacts to improve their feedback quality.

4.2 Affective factor

Feedback behaviors are mediated by students' imagined identities, as instantiated in their stances and motives (Yu & Hu, 2017b; Yu & Lee, 2015; Zhu & Mitchell, 2012). Lockhart and Ng (1995) discovered four stances (i.e. authoritative, interpretive, probing, and collaborative) in feedback responses. In this study, LEP students adopted probing and collaborative stances, which had an impact on the types and the amount of feedback they made. First, the probing stance was associated with students' motive to act as a sympathetic and patient reviewer, who took the initiative to check the author's intentions. That was why they made more clarification requests in content feedback, rather than dismissing the peer's writing as problematic. Second, the collaborative stance derived from students' motive to act as a helpful and constructive reviewer. Chan's self-imposed rule of at least one comment per page and Chow's meticulous endeavors to offer as many comments as possible demonstrated that the imagined identities led them to generate a large amount of feedback comments.

Conversely, HEP students tended to adopt authoritative and interpretative stances, which had an impact on the types of feedback they made. The authoritative stance was prominent in language feedback. HEP students judged whether the text was grammatical and made direct language changes on their peers' texts. To a certain extent, they acted as an authoritative judge, deciding the language quality and fixing the problems if deemed necessary. In content feedback, they operated in an interpretative stance, privileging their own responses over the writer's intentions (Lockhart & Ng, 1995). Yip's and Tam's interview extracts demonstrated that they foregrounded their feelings and impressions when reading the text. They focused on giving emotive responses, but did not (want to)

check the author's intentions. Therefore, they made fewer clarification requests in content feedback.

4.3 Sociocultural factor

The sociocultural context is the macro-contextual factor in peer feedback (Ellis, 2010). Previous studies have reported that students in the Asian contexts might refrain from making critical comments for fear of upsetting group harmony or threatening their peers' face (Allaei & Connor, 1990; Carson & Nelson, 1996). In addition, better English proficiency usually presupposes a more authoritative position. That was why Allen and Katayama (2016) found that students who perceived themselves to be less proficient than their peers refrained from making a large number of comments. However, in this study, both HEP and LEP students were able to make a substantial and similar amount of feedback. This does not mean that the sociocultural issue did not affect the students. In fact, in the interviews, LEP students confessed that they felt constrained by the face issue, but this issue was circumvented (if not solved) by the double-blind policy. Protected by anonymity, LEP students were allowed the freedom to offer critical and honest comments to their peers. Better still, they did not shy away from challenging the more proficient students, who might be otherwise presumed to hold an authoritative position if the identities were known. In this sense, anonymity can allow a safe space for LEP students, who are not compelled to hold back in peer feedback because of the face issue or the perceived proficiency gap (Huisman et al., 2017; Raes, Vanderhoven & Schellens, 2015).

4.4 Instructional factors

The instructional context is the micro-contextual factor in feedback (Ellis, 2010). In this study, LEP students reported that feedback quality was ensured partly because of two instructional reasons: feedback separation and time availability.

Although mixing content and language feedback does not affect the quality of revisions (Ashwell, 2000), previous studies have suggested to separate content issues from language issues when giving feedback (Allen, 2015; Yu & Lee, 2014). This study further demonstrated that such separation could help LEP students set a feedback focus and improve their feedback quality. As reported by the participants, it directed LEP students'

attention to either content or language issues so that they would not be overwhelmed or disoriented by treating every possible issue. It broke down the feedback task into two manageable sub-tasks, thereby reducing the task complexity for the LEP students and contributing to feedback quality (discussed later).

Time is an important factor in the composing process because students perform better if more time is allowed (Powers & Fowles, 1996). However, time has not been highlighted as a salient factor in feedback performance. Traditionally, peer feedback activities are held in a classroom and take up regular class time. As such, time pressure may limit students' willingness and/or ability to have deep engagement with their peers' writings and provide well-thought-out feedback (Chang, 2012). In this study, students were allowed ample time to make feedback at a pace they were comfortable with. They could check reference materials, think over difficult issues, and draft detailed comments. If being rushed through the task, LEP students might treat obvious and superficial issues, thereby limiting their feedback quality (more on this in the next subsection).

4.5 The interplay of factors and domains

Although the preceding subsections discuss each factor separately, this does not mean that feedback performance is influenced by a single factor. In fact, feedback performance is usually co-determined by an interplay of factors within and across the individual and contextual domains (Yu & Hu, 2017b; see also Figure 2). The relation between English proficiency and stance illustrates the interaction between the cognitive and affective factors within the individual domain. Recalling that LEP students made more suggestions and fewer direct changes in language feedback, we can follow that English proficiency provided a cognitive basis for the epistemic stance the students took. As they were less confident in English, they tended to be less assertive, while their more proficient/confident counterparts adopted an authoritative stance and made a larger amount of direct corrections. This confounding effect between cognition and affect lends further support to Allen and Katayama's (2016) observation that confidence in English proficiency influenced the types of feedback students provided.

Furthermore, feedback separation and time availability represent the interaction between the instructional and cognitive factors. In this study, separating content and language

feedback was a deliberate instructional design to scaffold LEP students. Researchers have pointed out that attention to linguistic elements might reduce the amount of cognitive resources available for generating or organizing content (Stevenson, Schoonen, & de Glopper, 2006; Yu & Lee, 2014). Feedback separation, therefore, guided LEP students to focus their attention on a particular aspect of the text, without risking cognitive overload. Additionally, as no time limit was set for the feedback task, LEP students were able to access reference materials as cognitive tools (Yoon, 2016) and have deep engagement with their peers' stories. These two instructional affordances reduced LEP students' cognitive pressure, and freed up more cognitive resources to perform the feedback task, thereby improving feedback quality.

5. Pedagogical Implications

Some EFL writing instructors might presume that lower English proficiency means poorer feedback performance. As the title of this paper asks, is this presumption a myth or a reality? The research results indicate that the distrust of LEP students' ability to make effective feedback is misplaced. LEP students do not differ in feedback amount or quality, but differ in feedback type. Therefore, we must reframe our question from *whether* to *how*: instead of questioning LEP students' ability to give feedback, we need to ask how we can engage and empower LEP students.

5.1 Engaging LEP students

LEP students are conventionally marginalized and conceptualized as beneficiaries rather than contributors in feedback dyads. However, studies have shown that LEP students can make valuable contribution and that HEP students can benefit from LEP students in peer feedback activities (Yu & Hu, 2017a; Yu & Lee, 2016b). Following this vein, this study further demonstrates that LEP students are able to give a similar amount of feedback with similar quality, compared with the feedback made by HEP students. Therefore, the role of LEP students as feedback contributors should be normalized and recognized in EFL writing classrooms. To this end, one possible pedagogical intervention is what Lee (2015) calls "intra-feedback." Intra-feedback requires reviewers of the same text to independently draft up comments and then meet together to discuss the strengths and weaknesses of each

other's feedback. Applying the concept of intra-feedback, EFL writing instructors can ask LEP and HEP students to review the same text and compare their feedback. The findings in this study can be points of comparison to be discussed in intra-feedback sessions. For instance, it is found that LEP students appear to be less assertive and more tolerant when giving feedback. Students can be prompted to explore why they give comments in divergent ways. What stances do they imply? Which feedback strategy is more helpful for writers to improve their writing? These questions enable students to understand how and why they comment differently, and sensitize them to different stances, projected and perceived, in their feedback behaviors. In this way, intra-feedback activities can engage LEP students and bring out their dual contribution: (a) student writers can benefit from LEP students' feedback; and (b) fellow feedback givers can learn from LEP students through discussion of varied stances and strategies, thereby improving their feedback-making repertoires.

5.2 Empowering LEP students

The second implication encourages EFL writing instructors to create an empowering environment for LEP students. As peer feedback is a site for collaborative learning (Lee, 2017), instructors need to make effort to unleash the potential of LEP students. To start with, instructors must realize that English proficiency is one of the many mediating factors in feedback activities. The conceptual framework discussed in Section 4 can provide a heuristic for instructors to recognize the impacts of cognitive, affective, sociocultural, and instructional factors on feedback performances. Additionally, the sociocultural theory of learning believes that the quality and quantity of scaffolding differ across proficiency levels (Lantolf, 2012). For instance, HEP students can rely on their English proficiency to achieve feedback quality, but LEP students may need extra mediational tools (i.e. corpora and dictionaries mentioned by the participants) to fulfil the task. Bitchener and Knoch (2009) aptly point out that feedback should be provided based on students' readiness. As the levels of readiness vary across students, instructors need to carefully orchestrate mediating affordances in a feedback task to empower LEP students. In this study, anonymity, access to technology and information artifacts, use of L1, feedback separation, availability of ample time were all important mediational means to empower LEP students, who might

otherwise be straightjacketed if the English language was the only resource allowed in the feedback task.

6. Conclusion

This study contributes to a growing body of research in understanding feedback performance of LEP students. It is found that HEP and LEP students do not differ in feedback amount or feedback quality, but differ in feedback type. LEP students tend to make more clarification requests in content feedback, more suggestions and fewer direct changes in language feedback. Based on students' interviews, a conceptual framework is proposed to explain how feedback performance is enabled or constrained by cognitive, affective, sociocultural, and instructional factors. This framework can be a heuristic for EFL writing instructors to create facilitative conditions for LEP students. This study and a growing number of others (e.g. Yu & Hu, 2017a, b; Yu & Lee, 2016b) attempt to demonstrate the importance of reframing the question from *whether* to *how*. Instead of doubting LEP students' ability to provide effective feedback, we are obligated to think about how to engage and empower LEP students.

One limitation of this study is its relatively small sample size (23 HEP students and 23 LEP students). There is a need to conduct replication studies to ascertain whether the statistically non-significant results in feedback amount and quality are borne out with a larger sample size. Additionally, this study is limited by the proficiency gap of the participants. As Allen and Katayama (2016) note, "lower" English proficiency is a relative rather than an absolute concept. This study compared students at the CEFR B2 and B1 levels because they represented the lower and higher ends of an intact cohort in a naturalistic feedback activity. Although a mixed-methods design has been adopted to triangulate quantitative and qualitative results, caution needs to be exercised to generalize the findings to other populations. For instance, when students at CEFR C1 and B1 levels are compared, would the two groups display significant differences in feedback performance? Is there a proficiency threshold for students to identify and actualize instructional and sociocultural affordances, thereby empowering themselves in feedback tasks? The answers to these questions would shed important light on the role of LEP

students and what EFL instructors can do to maximize the learning potential in peer feedback.

Appendix A

An example worksheet for the peer feedback task.

1. Your four digit code is 7634. Create a folder named 7634.
2. You are going to comment on Story #1487. Locate the story from the database in the course management system. Copy it twice into your own folder (7634).
3. In your folder, rename the two Word files as follows: #1487-7634-content and #1487-7634-language.
4. Comment on the content first and then the language. Make content feedback in the file #1487-7634-content, and language feedback in the file #1487-7634-language. Use Track Changes and marginal comments to make feedback.
5. There is no time limit for this task. You are encouraged to use the Internet access to consult reference materials.

Appendix B

Interview questions for Tsai.

1. Please take a look at the feedback you made. Do you spot particular patterns in your feedback? Why do you comment in these patterns?
2. Do you recall particular moments in the feedback task when you feel empowered or impeded? Why?
3. Do you think your feedback pattern is different from your peers'? If so, how and why?
4. Do you think you did a good job in giving feedback? What factors have helped or inhibited you in giving quality feedback?
5. In this feedback task, you made many feedback comments, even slightly more than your peer did. Why?
6. I notice that you made quite a few clarification requests in the content feedback. Why?
7. In the language feedback, you made quite many marginal comments but few direct changes. Why?
8. Anything else you want to share about the feedback task?

References

- Allaei, S., & Conner, U. (1990). Exploring the dynamics of cross-cultural collaboration in writing classrooms. *The Writing Instructor*, 10, 19–28.
- Allen, D. (2015). Personal and procedural factors in peer feedback: A survey study. *Komaba Journal of English Education*, 6, 47–65.
- Allen, D., & Katayama, A. (2016). Relative second language proficiency and the giving and receiving of written peer feedback. *System*, 56, 96–106.
- Allen, D., & Mills, A. (2013). Peer feedback in the academic English classroom: A pilot study. *Komaba Journal of English Education*, 4, 27–52.
- Allen, D., & Mills, A. (2016). The impact of second language proficiency in dyadic peer feedback. *Language Teaching Research*, 20(4), 498–513.
- Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method?. *Journal of Second Language Writing*, 9(3), 227–257.
- Bitchener, J., & Knoch, U. (2009). The value of a focused approach to written corrective feedback. *ELT Journal*, 63(3), 204–211.
- Carson, J. G., & Nelson, G. L. (1996). Chinese students' perceptions of ESL peer response group interaction. *Journal of Second Language Writing*, 5(1), 1–19.
- Caulk, N. (1994). Comparing teacher and student responses to written work. *TESOL Quarterly*, 28, 181–188.
- Chang, C-F. (2012). Peer review via three modes in an EFL writing course. *Computers and Composition*, 29, 63–78.
- Chang, C. (2015). Teacher modeling on EFL reviewers' audience-aware feedback and affectivity in L2 peer review. *Assessing Writing*, 25, 2–21.
- Diab, N. M. (2010). Effects of peer-versus self-editing on students' revision of language errors in revised drafts. *System*, 38(1), 85–95.
- Ellis, R. (2010). Epilogue: A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, 32(2), 335–349.
- Glaser, B. (1992). *Basics of grounded theory analysis: Emergence vs forcing*. Mill Valley, CA: Sociology Press.

- Huisman, B., Saab, N., van Driel, J., & van den Broek, P. (2017). Peer feedback on college students' writing: exploring the relation between students' ability match, feedback quality and essay performance. *Higher Education Research & Development*, 36, 1433–1447.
- Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83–101.
- Kamimura, T. (2006). Effects of peer feedback on EFL student writers at different levels of English proficiency: A Japanese context. *TESL Canada Journal*, 23(2), 12–39.
- Lantolf, J. P. (2012). Sociocultural theory: A dialectical approach to L2 research. In S. M. Gass & A. Mackey (eds.). *The Routledge handbook of second language acquisition* (pp. 57–72), Routledge.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Routledge.
- Lee, M. K. (2015). Peer feedback in second language writing: Investigating junior secondary students' perspectives on inter-feedback and intra-feedback. *System*, 55, 1–10.
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer.
- Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language Learning*, 45(4), 605–651.
- Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18(1), 30–43.
- Min, H. T. (2005). Training students to become successful peer reviewers. *System*, 33(2), 293–308.
- Min, H. T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15(2), 118–141.
- Nelson, G. L., & Murphy, J. M. (1993). Peer response groups: do L2 writers use peer comments in revising their drafts? *TESOL Quarterly*, 27, 135-142.
- Oxford University. (2001). *Quick Placement Test* (Paper and Pen User Manual). Oxford: Oxford University Press.

- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Powers, D. E., & Fowles, M. E. (1996). Effects of applying different time limits to a proposed GRE writing test. *Journal of Educational Measurement*, *33*, 433–452.
- Raes, A., Vanderhoven, E., & Schellens, T. (2015). Increasing anonymity in peer assessment by using classroom response technology within face-to-face higher education. *Studies in Higher Education*, *40*(1), 178–193.
- 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.
- Stevenson, M., Schoonen, R., & de Glopper, K. (2006). Revising in two languages: a multi-dimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing*, *15*, 201–233.
- Tsui, A. B., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments?. *Journal of Second Language Writing*, *9*(2), 147–170.
- Tsui, A. B., & Ng, M. (2010). Cultural contexts and situated possibilities in the teaching of second language writing. *Journal of Teacher Education*, *61*(4), 364–375.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, *15*(3), 179–200.
- Yoon, C. (2016). Concordancers and dictionaries as problem-solving tools for ESL academic writing. *Language Learning & Technology*, *20*(1), 209–229.
- Yu, S., & Lee, I. (2014). An analysis of Chinese EFL students' use of first and second language in peer feedback of L2 writing. *System*, *47*, 28–38.
- Yu, S., & Lee, I. (2015). Understanding EFL students' participation in group peer feedback of L2 writing: A case study from an activity theory perspective. *Language Teaching Research*, *19*, 572–593.
- Yu, S., & Lee, I. (2016a). Peer feedback in second language writing (2005–2014). *Language Teaching*, *49*(4), 461–493.

- Yu, S., & Lee, I. (2016b). Understanding the role of learners with low English language proficiency in peer feedback of second language writing. *TESOL Quarterly*, 50(2), 483–494.
- Yu, S., & Hu, G. (2017a). Can higher-proficiency L2 learners benefit from working with lower-proficiency partners in peer feedback?. *Teaching in Higher Education*, 22(2), 178–192.
- Yu, S., & Hu, G. (2017b). Understanding university students' peer feedback practices in EFL writing: Insights from a case study. *Assessing Writing*. 33, 25–35.
- Zhao, H. (2011). Using learners' diaries to investigate the influence of students' English language proficiency on peer assessment. *Journal of Academic Writing*, 1(1), 126–134.
- Zhu, W., & Mitchell, D. A. (2012). Participation in peer response as activity: An examination of peer response stances from an activity theory perspective. *TESOL Quarterly*, 46(2), 362–386.

