



The bilateral benefits of providing and receiving peer feedback in academic writing across varying L2 proficiency

Ying Gao^{a,*,1}, Qi An^{b,**}, Christian D. Schunn^{c,***,2}

^a Ministry of Education (MOE) Language Training Center and School of Foreign Languages, Northeast Normal University, Changchun, China

^b English Faculty, Dalian Hi-Tech Industrial Zone No.1 Middle School, Dalian, China

^c Learning Research & Development Center, University of Pittsburgh, Pittsburgh, USA

ARTICLE INFO

Keywords:

Peer review
Providing feedback
Receiving feedback
Bilateral benefit
L2 proficiency
Writing evaluation

ABSTRACT

Although the effectiveness of peer review has been examined, few have tested the joint benefits of providing and receiving feedback features and quality in L2 contexts. The present study investigated variation in key features and quality of feedback provided and received by high and low L2 proficiency students and its benefits on revision in the authentic setting of students experiencing both roles of providing and receiving peer feedback in English academic writing. Analyses on two drafts from 50 students, 1356 idea units of anonymous implementable peer feedback and back-evaluation ratings revealed that the combined effect of providing and receiving feedback on revision was prominent for all students. But the bilateral benefits of providing and receiving feedback features were more significant for high proficiency students, while providing feedback features and quality benefited low proficiency students more. The study implies that students need more bilateral training with both providing and receiving feedback.

1. Introduction

Peer review/feedback processes (consisting of qualitative feedback and quantitative assessment) involve the activities undertaken by learners to consider and specify the level, value, or quality of a product (Topping, 2009: 20), and to obtain, understand and use feedback (Winstone et al., 2022: 224) within a process writing approach. Although it has been argued to be an effective and efficient strategy for teaching and learning in both first language (L1) and second language (L2) writing contexts in recent decades (Topping, 1998; Hyland & Hyland, 2006; Nicol et al., 2014; Double et al., 2020; Chang et al., 2021), doubts on the helpfulness of peer review remain (Allen & Katayama, 2016; Wu, 2019). In particular, the joint benefits of providing and receiving peer feedback on learning is largely unknown as previous studies have mainly focused on the effectiveness of receiving feedback on author draft revision (Gao, Schunn, & Yu, 2019; Patchan & Schunn, 2016), or the benefits of providing feedback to peers (Cho & Cho, 2011;

Patchan & Schunn, 2015; Yu, 2019). In the limited number of studies that involved both providing and receiving feedback, the concerns were on comparing the contributions of providing with receiving feedback (Ludstrom & Baker, 2009; Li et al., 2012; Ion et al., 2018) or learner perception of the benefits through qualitatively inquiries (Cao et al., 2019) rather than looking at the joint effect of providing and receiving feedback in the bilateral process of peer review (Wu & Schunn, 2020). In real school settings, peer review involves both providing and receiving feedback. Not only is it likely that both providing and receiving feedback contribute to learning improvement (Li et al., 2012), experimental studies that only allow one kind of learning opportunity are highly artificial, and correlational investigations of one component need to control for variation in the other component. Therefore, it is important to explore how providing and receiving feedback can jointly contribute to learning in authentic combination rather than to only look at the benefits in isolation.

Furthermore, there is a general call for more complex, authentic

* Correspondence to: MOE Language Training Center & School of Foreign Languages, Northeast Normal University, Room 327; 5268 Renmin Street, Changchun, Jilin 130024, China.

** Correspondence to: English Faculty, Dalian Hi-Tech Industrial Zone No.1 Middle School, Room 164; 221 Xiaopingdao Street, Dalian, Liaoning 116023, China.

*** Correspondence to: Learning Research and Development Center, University of Pittsburgh, Room 424B; 3420 Forbes Ave, Pittsburgh, PA 15260, USA.

E-mail addresses: gaoy687@nenu.edu.cn (Y. Gao), ann2738@qq.com (Q. An), schunn@pitt.edu (C.D. Schunn).

¹ orcid.org/0000-0003-3565-2351

² orcid.org/0000-0003-3589-297X

tasks in higher education (e.g., project-based learning in mathematics, engineering, and science) and requirements for graduate students to produce publication-quality products in order to graduate. Writing is a critical aspect of all such complex tasks, both from an evaluation perspective and from a writing-to-learn perspective. Peer evaluation is a particularly relevant pedagogy because it assists in making feedback timely (in comparison to instructor feedback) (Wu & Schunn, 2020), of greater quantity (when implemented as multi-peer feedback) (Cho & Schunn, 2007), and because students benefit from providing feedback to their peers (Cho and MacArthur, 2011; Berggren, 2015). However, when considering L2 students, many concerns have been raised, particularly for more complex forms of writing. Extensive research has shown that L2 students do benefit from peer feedback, but most of those studies have focused on very simple, introductory forms of writing that is typical in early EFL coursework (Allen & Mills, 2016; Chong, 2017). Much more research needs to be done on L2 learners engaged in peer feedback in more complex writing, particularly because many non-English speaking countries are requiring graduate education to be conducted in English or work products to be produced in English. And such writing is particularly challenging for L2 writers as they have to struggle with an extra load related to language issues (Gao, Schunn, & Yu, 2019).

Although peer feedback activities may provide benefits for students of different L2 proficiency, what is actually 'provided' and 'received' may differ considerably (Allen & Mills, 2016; Yu & Lee, 2016b) in terms of whether the feedback contains the features that guide a successful revision (e.g., identifying the problem and providing possible solutions) or whether the feedback was helpful. L2 proficiency is a significant factor influencing the effects of peer review (Hyland & Hyland, 2006) as it can influence both writing quality and feedback quality. Unfortunately, feedback quality is inadequately considered and rarely combined with feedback features in previous studies, and the impact of L2 proficiency on student benefits in peer review has been relatively under-researched as most studies have tried to control the proficiency variable and focused on student contributions in peer review (Allen & Katayama, 2016; Allen & Mills, 2016).

Therefore, this quasi-experimental design study is conducted in a setting where student roles of both providing and receiving peer feedback are included, and it focuses on the question: what kinds of peer feedback is particularly useful for L2 learners working on research papers? Does proficiency in English change the kind of peer feedback experiences that are particularly helpful?

2. Literature review

2.1. Learner benefits from providing and receiving feedback

According to the Zone of Proximal Development (ZPD) in sociocultural theory (Lantolf & Thorne, 2006), peer review enables students to transit from other-regulation (facilitation by peers of varying abilities) to self-regulation, building upon not only on the feedback received but also on the feedback provided and the interaction among peers within their respective ZPDs (Ion et al., 2018). Peer review enables positive interdependence, equal participation, individual accountability and group processing as advocated in cooperative learning theory (Slavin, 1990). Consequently, in the past decades, peer review has been widely used in L2 writing contexts and has yielded diverse and mixed effects on many aspects of learning. On the one hand, peer review is believed to contribute to students' cognitive, affective, social-cultural, and linguistic growth (Liu & Hansen, 2002). For example, it fosters improved writing performance, higher-order cognitive abilities, writing autonomy, and audience awareness as "cognitive and meta-cognitive benefits can accrue before, during, or after the peer assessment" (Topping, 2009: 23), and the complexity of the reviewing process nurtures the development of critical thinking skills and problem-solving skills, as well as the development of writer autonomy, audience awareness and multiple linguistic skills (Gao, Schunn, & Yu, 2019; Min, 2008). On the other

hand, however, doubts about the effectiveness of peer review remain as peers are often skeptical about the authority of the received feedback and some students perceive providing feedback as a waste of time (Allen & Mills, 2016; Yu & Lee, 2016b). In providing feedback, students tend to focus on local form issues rather than global content issues as they perceive peer review as "a task to correct grammar and format rather than an opportunity to exchange and discover ideas with the writers" (Min, 2008: 300). As a result, the less constructive feedback students receive negatively influences their ability to select valid suggestions when revising their own writing (Patchan & Schunn, 2015, 2016). Additionally, students from teacher-centered cultures see the teacher as the only source of authority and demonstrate strong reliance on teacher feedback (Allen & Mills, 2016), and they are negative towards peer review because they lack confidence, social and communicative skills, and the relevant language knowledge needed in such activities (Zhang, 2018). In practice, however, feedback is not just a descriptive paradigm, but also a constructive one; students learn by forming meaning from receiving feedback, and by constructing meaning themselves (Nicol et al., 2014).

To compare the relative benefits of providing and receiving feedback for ESL students, Lundstrom and Baker (2009) conducted a highly controlled experimental study with two conditions: Receivers only received feedback without providing any comments whereas the Providers were limited to providing feedback without receiving feedback. Providers produced higher quality writings than Receivers, and lower ability Providers benefited more than higher ability Providers did. Cho and MacArthur (2011) similarly found that students in the reviewing condition significantly outperformed those in the reading and no-treatment control conditions in the quality of subsequent writings, and that providing 'problem detection' and 'solution and suggestion' were positively correlated with writing outcomes for students in the reviewing condition. These studies suggested that providing feedback helped students improve their learning more than receiving feedback did (Berggren, 2015). Nevertheless, the experimental design of assigning learners only one role is "short of ethical consideration and may break the peer feedback continuum" (Cao et al., 2019: 103), and the deliberate separation of the two intertwined roles overshadows the interactive nature of peer review and thus lacks ecological validity (Yu & Lee, 2016a). Students do not just benefit from providing critical feedback and explanations but also from receiving feedback through which they get to know deficiencies in their work and interpret readers' needs (Nicol et al., 2014).

A few other studies supporting the learning-by-reviewing hypothesis were carried out in real situations with peers both providing and receiving feedback (Trautmann, 2006; Cho & Cho, 2011; Berggren, 2015). Cho and Cho (2011) compared the effects of providing and receiving feedback on reviewers' revised drafts and found that receiving feedback was less powerful in enhancing writing performance than providing feedback, and providing microscope-level criticism and praises appeared especially useful to learning. Berggren (2015) found that ESL learners raised their audience awareness through providing feedback and were motivated to revise their own drafts by the content reviewed, thereby improving writing skills. However, in Trautmann (2006), 70 % of learners believed they revised because of others' feedback rather than self-provided feedback. Being both a feedback provider and receiver were found to be mutually beneficial in Yu and Wu (2016), but their focus was only on feedback quality and the study was conducted on the task of question-generation rather than writing. Recent studies revealed that learners perceived different patterns of learning from peer feedback: giving and receiving, receiving only, neither giving or receiving, and giving only (Cao et al., 2019), and both amount of received and provided feedback were associated with being more likely to make revisions (Wu & Schunn, 2020). The inconsistent findings necessitate more empirical research to investigate learning in providing and receiving feedback and the factors that may shape their learning process and outcome (Yu, 2019).

Additionally, investigation of the effectiveness of peer feedback often draws attention to the heterogeneous contents of peer feedback, in particular the feedback features like explicitly identifying a problem and giving suggestions or solutions (Nelson & Schunn, 2009; Wu & Schunn, 2020), whereas the quality of feedback has been scarcely examined. Good quality peer feedback can develop critical reflection and encourage more giving, receiving and acting on feedback (Liu and Carless, 2006). Although not many studies have focused on this aspect of peer review, it is generally agreed that there are many approaches to measuring peer feedback quality, for example, the accuracy of the comments in addressing problems (Gao, Schunn, & Yu, 2019), the number and length of comments in providing sufficient persuasion (Patchan, Schunn, & Clark, 2018), the form of comments (including aspects and functions) like problem identification and constructive suggestion/solution (Huisman, Saab, Driel, & Broek, 2017; Li, Liu, & Steckelberg, 2010; Nelson & Schunn, 2009; Yu & Wu, 2016), as well as the effect of feedback on document revision (Cho & Cho, 2011; Wu, 2019). A more comprehensive approach proposed recently is to take a recipient-centric approach (followed in the present study) by finding out how recipients judge the helpfulness of feedback which is usually a holistic multidimensional judgment (Zong, Schunn, & Wang, 2020).

2.2. Learner L2 proficiency in peer review

Although the significance of L2 proficiency in determining students' ability to provide and utilize peer feedback is well recognized, its impact on peer feedback has been under-explored (Yu & Lee, 2016b). The limited body of existing research mainly fall into two strands. The first strand focuses on the contributions of students with different L2 proficiency in peer review. Quantitative methods are usually used to compare how high proficiency students (HPs) and low proficiency students (LPs) differ in their feedback products, and qualitative designs are used to describe the dynamics in mixed-proficiency feedback groups (Wu, 2019). In general, HPs could provide more suggestions on their peer's writing than students with LP, and LPs tend to be portrayed as less able to contribute to peer feedback, and that they primarily play the role of a receiver rather than a giver (Allen & Mills, 2016). However, Yu & Lee (2016b) found in their qualitative study that LPs are able to contribute to peer feedback by providing a range of comments on various aspects of writing and help their peers enhance their writing quality. Besides, findings that HPs and LPs did not differ in feedback quality suggest that HPs and LPs were able to assume the dual roles of contributors and beneficiaries in feedback activities (Yu & Lee, 2016b; Wu, 2019).

The second strand of research concentrates on the benefits of peer review on students of different L2 proficiency. This line of research is even less in number and usually conducted in comparing the effects of L2 proficiency matching. It is found that the effect of high proficiency reviewers providing more comments than low reviewers was particularly salient in asymmetrical matching groups (e.g., High \times Low or Low \times High). And high proficiency students could learn from the feedback given by low proficiency students, and develop their audience awareness and L2 writing knowledge, and high proficiency students can benefit by receiving feedback from high proficiency students and high proficiency students can learn from giving feedback to high proficiency students (Allen & Mills, 2016; Lundstrom & Baker, 2009).

Additionally, despite of the abundant findings on the benefits of peer review on L2 learners, previous studies have focused heavily on simple and short writing tasks (Allen & Mills, 2016; Chong, 2017). Authentic academic writing typically is much longer and also involves mastering specific genres (Swales, 2004; Kwan, 2006). Relatively little is known about how peer review behaviors and outcomes vary as the genre is varied or becomes more complex. A literature review in particular is an authentic and especially complex genre that balances content (e.g., claims and evidence; rhetorical moves in establishing the larger argument), higher-level aspects of language (e.g., flow and organization of complex ideas across long segments of text), and lower-level aspects of

language (e.g., spelling, grammar, and word choice) (Gao, Schunn, & Yu, 2019; Kwan, 2006). Relative to other forms of academic writing genres, the hallmark move structures in literature reviews (e.g., the three-move structure which covers (1) establishing the territory of one's research, (2) creating a research niche, (3) occupying the research niche) creates many challenges for learners and might influence learner benefits through peer review. For example, one past study of peer reviewing of literature reviews by L2 students found that they often struggled to notice certain move structure problems or provide useful suggestions for peers (Gao, Schunn, & Yu, 2019).

In sum, learner L2 proficiency may determine students' contribution to peer feedback as well as their benefits from it. However, due to the paucity of relevant studies, the issue of how providing and receiving feedback features and quality jointly contribute to revision improvement, and how they are related to learners with different L2 proficiency remain unclear, especially in complex writing tasks.

3. The current study

The current study aims to investigate the joint impact of providing and receiving feedback on different L2 proficiency students, focusing on the bilateral impact of peer review on draft revision. The study was conducted in the context of postgraduate students writing literature reviews. The study was designed to answer two research questions:

- (1) What kinds of provided and received peer feedback is particularly useful for L2 students working on academic writings in terms of key feedback features and feedback quality?
- (2) Does proficiency in English change the kind of provided and received peer feedback experiences that are particularly helpful to draft revision in academic writing?

3.1. Research setting and participants

The study was conducted within an 'Intercultural Communication' course for first year students in an MA program in Linguistics at a public research-oriented university in China. As part of the course, students were required to write two drafts of a literature review (2000 words) in English and conduct two rounds of peer review for their course thesis. The assignment was given at the beginning of the semester, leaving adequate time for students to read the literature and write the first draft.

To facilitate both writing and reviewing, a rubric (See Appendix) was designed by the instructor, providing detailed task requirements and reviewing standards. Following Gao, Schunn, & Yu (2019), the rubric was adapted on basis of the adapted CARS (create a research space) model (Swales, 1990, 2004), the three-move structure of literature review writing proposed by Kwan (2006), and expert input. It consisted of three dimensions: Introduction, Body paragraphs and Conclusion, each containing specific descriptions and question prompts (e.g., (1) Write your ideas on how this paper works in presenting the central topic and establishing its importance. (2) Did the paper clarify the argument involvement in the literature? Did the paper clarify the relationships of the arguments and counter arguments?). Within all three dimensions, students were asked to pay attention to clarity of writing, logic and organization, effectiveness/persuasiveness of writing, and English writing conventions.

Participants were native speakers of Chinese and had learned English for over thirteen years. Through both undergraduate and postgraduate writing courses, students had been informed about general conventions in academic writing, and they had received instruction on writing literature reviews. All 50 students in the course completed the writing and reviewing assignments, and therefore, all were included in data analyses. The L2 proficiency of the participants was determined by their scores on the Test of English Majors Band 8 (TEM 8, an annual national

English proficiency test for English majors in China, with a possible total of 100 points). Using a median split method, which is often used in this research area and that extreme group designs (selecting only very high and very low) are criticized for offering worse generalizability, students were divided into 25 high proficiency students (HPs) ($M = 71.3$; $SD = 5.4$) and 25 low proficiency students (LPs) ($M = 60.6$; $SD = 3.0$) on basis of their TEM 8 scores, producing a statistically significant difference between the two groups ($t(48) = 8.59$, $p < .001$). In general, the L2 proficiency of the students was approximately between 78 and 112 on the Test of English as a Foreign Language, which corresponds to the “intermediate” and “advanced” categories respectively.

3.2. Procedure

Using *Peerceptiv* (an online peer review writing system, <https://asia.peerceptiv.com>), each student reviewed three randomly-assigned peers' drafts in a double-blinded manner. Online multi-peer anonymous peer review is ideal in this designed study as “cumulative effect from peers (regardless of level) constitutes more favorable learning conditions” (Gao, Schunn, & Yu, 2019), and the anonymity of provider and receiver identity guarantees peer comments to be based solely on the text itself and the incorporation of peer feedback to be determined by the feedback quality itself, rather than personal relationships (Cote, 2014). Further, different from oral peer feedback or face-to-face peer written feedback, the asynchronous feature of the system is most favorable for complex writing tasks as in this case of writing literature reviews as it allows more time for learners to reflect on the content of writing rather than just focusing on superficial errors. After completing their first drafts, participants uploaded them to *Peerceptiv* (Week 15) and spent 4 weeks reviewing peers' work and revising their own drafts after receiving peer feedback, and another 2 weeks giving back-evaluation comments and ratings on the overall helpfulness (quality) of feedback they received.

Students were trained on how to give helpful comments through comparative illustrations of good and bad examples in class and video clips demonstrating a checklist of DOs and DON'Ts in reviewing before and during the process of peer review. To encourage engagement, students were graded for course points on the accuracy and helpfulness of the feedback they provided. Learners uploaded their revised drafts to the system as the final course paper (Week 21).

3.3. Measures

3.3.1. Text quality

Text quality was based upon mean expert ratings. For both Draft 1 and Draft 2, two writing experts with instructional experience rated the equally weighted three dimensions of the rubric (see [Supplemental Appendix A](#) in the online version of the journal) on 7-point Likert scales (later converted to a percentage), running from “difficult to read, poor, fair, OK, good, very good, to excellent”, and the expert ratings had high inter-rater reliability ($K = 0.80$ for Draft 1; $K = 0.82$ for Draft 2). T-tests confirmed a significant growth from Draft 1 to Draft 2 and significant differences between high and low proficiency groups within each draft (see [Table 2](#)).

3.3.2. Feedback quality

Feedback quality was conceptualized holistically in terms of the feedback recipient's perceptions (Zong, Schunn, & Wang, 2020) and was based upon the back-evaluation ratings from the recipients on how helpful the feedback they received was to the improvement of their writings by rating on a 5-point Likert scale per dimension per review (5 - most helpful, 4 - very helpful, 3 - helpful, 2 - not very helpful, 1 - unhelpful). Every student rated the three peers who provided them feedback and every student was rated by three peers who received their feedback. Back-evaluation scores were averaged across the three dimensions and three reviews to produce measures of the average quality of reviews each student provided and received. As students were found

not willing to giving back-evaluations (Nelson & Schunn, 2009), a special caution was taken to constantly remind them to do the task, consequently, all completed the back-evaluation step.

3.4. Feedback coding

Feedback was segmented into idea units before the data was content coded. An idea unit refers to a feedback segment expressing a single topic, either in the form of phrases or sentences (Nelson & Schunn, 2009). Two independent coders (the same two draft raters) segmented all the peer feedback into idea units, producing a total of 2454 idea units.

Idea units that only contained praise or a summary of the document (1098 idea units) were excluded from further analysis because they were not implementable ideas for revision. The presence or absence of two key feedback features were analyzed within the implementable comments: *Identifying problems* and *Giving suggestions/solutions* (shortened as *Problem and Suggestions/solutions*) as they have been reported to be the most significant features predicting revision in previous studies and they appeared most frequently in the data. Following Nelson & Schunn, 2009, each of these two features of feedback is defined and an example is provided from the data (See [Table 1](#)).

Inter-rater Kappa values for both feedback features between the two coders were above 0.8, indicating high inter-rater reliability. To minimize coding inconsistency, all comments were double-coded, and disagreements were discussed and resolved.

3.5. Data description and analysis

In total, our data include two drafts of the 50 documents, expert ratings of the documents (text quality), 1356 idea units coded in terms of two key features of implementable comments (*Problems* and *Suggestions/solutions*), and back-evaluation ratings (feedback quality) of all received and provided reviews. After revision, second drafts were longer and better in quality, providing a necessary foundation for exploring improvements on the basis of peer feedback provided and received (See [Table 2](#)).

To answer the research questions, a series of multiple linear regression tests were run to test the predictive strength of the amounts of features and quality of provided only (1), received only (2), and provided + received feedback (3) on the improvement of revision in Draft 2 quality overall (RQ1) and separately for both HPs and LPs (RQ2). Three types of models were run: 1) a simple partial correlation of just the isolated feedback feature with Draft 2 score, 2) a stepwise forward regression that sequentially added predictors when their partial correlations were statistically significant, and 3) the full multiple regression that included all predictors. In each case, the models focus on Draft 2 score controlling for Draft 1 score rather than calculated change scores (Draft 2 minus Draft 1) because change scores have regression-to-the-mean problems that produce spurious correlations when they are used

Table 1
Types and definitions of implementable comments along with an example comment.

Type	Definition	Example
<i>Problems</i>	Giving an explicit diagnosis of what the problem is	<i>The writer defines all the elements that were connected with the title but barely bonded them together, which makes this article confusing.</i>
<i>Suggestions/solutions</i>	Offering (1) general advice on the writing content and language problems or (2) specific ways to solve the problem, and more detailed methods on how to modify or improve draft quality	<i>(1) Except introducing the main topic, the author should provide the background information of this study. (2) For instance, the writer can give some detailed information about management strategy.</i>

Table 2
Mean and standard deviations for draft quality (rating in %) along with t-test values from the comparison across drafts and proficiency group.

Learners	N	Draft 1			Draft 2		
		Mean	SD	t	Mean	SD	t
All students	50	70.5%	6.4%		77.0%	7.2%	11.1***
High proficiency students	25	73.0%	5.5%	2.7*	79.0%	7.9%	2.1*
Low proficiency students	25	68.1%	6.4%		75%	5.9%	9.5***

Notes. * = $p < .05$, ** = $p < .01$, *** = $p < .001$

as the dependent variable in multiple-regression. Collinearity was checked by examining correlations among feedback features and Variance Inflation Factors (VIF). In addition, to answer RQ2, frequency of each feedback feature for each review and the feedback quality for HPs vs. LPs were compared through t-test and a series of 2×2 between-subjects ANOVAs were run to explore the main effects of the interaction between author and reviewer proficiency on the frequency of comments with each feedback feature and overall feedback quality.

4. Results

4.1. The benefits of providing and receiving feedback features and quality on revision (RQ1)

As students performed double roles in this peer review activity, revision is assumed to involve the mutual effects of both providing and receiving feedback. Thus, the benefits of providing and receiving feedback on draft revision were considered jointly. Specifically, the goal was to determine whether all features and quality of feedback jointly predicted revision or just certain types did.

Although the total amount of provided feedback and received feedback was the same, the feedback provided and received by each student varied. Therefore, taking students as authors who received feedback and as reviewers who provided feedback produced two different sets of data for analysis. Peers very commonly identified many *Problems* ($M = 13.9$, $SD = 7.7$), and in almost as many cases, *Suggestions/Solutions* ($M = 13.2$, $SD = 8.2$) were offered. The quality of peer feedback ($M = 4.2$, $SD = 0.7$) was above 4/5 (80%) overall, indicating a very positive helpfulness rating by learners.

As shown in Table 2, t-test revealed significant improvement ($t=11.1***$) in Draft 2 ratings ($M = 77.0$; $SD = 7.2$) as compared with that of Draft 1 ($M = 70.5$; $SD = 6.4$), with an improvement rate of 9.2 %.

Multiple linear regression analyses demonstrated the best linear combination of providing and receiving feedback in predicting improvement. Within the stepwise model and the full model, three tests were run to compare the predictive strength of providing only (1), receiving only (2) and the joint strength of providing + receiving feedback (3).

Table 3

Across the three tested models, the strength (partial r / standardized β) of the relationship of each providing and receiving comments of each type to draft 2 quality, controlling for Draft 1.

Predictors	Partial r with Draft 2	Regression β Predicting Draft 2					
		Stepwise model			Full model		
		Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
Provided (n = 50)							
<i>Problems</i>	0.27*	-	-	-	0.10	-	0.12 [†]
<i>Suggestions/Solutions</i>	0.60***	0.38***	-	0.32***	0.29**	-	0.26***
<i>Feedback quality</i>	0.35**	-	-	0.13*	0.10	-	0.11 [†]
Received (n = 50)							
<i>Problems</i>	-0.04	-	-	-	-	0.05	0.10
<i>Suggestions/Solutions</i>	0.11	-	0.25**	0.24***	-	0.21*	0.20**
<i>Feedback quality</i>	-0.07	-	-	-	-	-0.09	-0.02
R ²	NA	0.81	0.74	0.87	0.83	0.75	0.89
Max VIF	NA	1.12	1.03	1.23	1.70	1.47	1.73

Notes. [†] = $p < .1$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Results shown in Table 3 clearly reveals that within both the stepwise model and the full model, R² is highest in Test 3, indicating that the joint strength of providing and receiving peer feedback outweighs that of either providing (Test 1) or receiving (Test 2) feedback alone. However, although not shown in the table, it is important to note that Draft 1 quality was a strong predictor of Draft 2 quality in all tests with both the stepwise model (Test 3: $\beta = 0.74$, $p < .001$; meaning that 55 % (0.74*.74) of the variance in Draft 2 was explained by Draft 1) and the full model (Test 3: $\beta = 0.78$, $p < .001$), as expected. Therefore, the features and quality of providing and receiving feedback strongly predicted revision improvement, accounting for 76 % (i.e., (89–55 %) / (1 – 55 %)) of change in Draft 2 rating.

The two feedback features in the provided and received comments predicted growth in Draft 2 quality. In particular, providing and receiving *Suggestions/Solutions* significantly predicted improvements in revision across three models and six tests, except for receiving with partial correlation. Providing *Problems* and feedback quality were significant indicators of Draft 2 changes in two and three models respectively. A closer observation of the tests reveals that, providing *Problems* contributed to draft improvement only in test 3 of the full model, and the quality of provided feedback predicted Draft 2 improvement only in test 3 of the stepwise model and the full model. That is to say, providing *Problems* and feedback quality significantly contributed to draft change only when both providing and receiving feedback were performed.

Overall, providing and receiving feedback jointly contributed to revision more than either providing or receiving peer feedback. In terms of relative contributions, the quantity of feedback features (especially *Suggestions/Solutions*) was more beneficial than that of feedback quality, and providing feedback was more predictive to draft changes than that of receiving feedback.

4.2. The helpfulness of providing and receiving feedback on draft revision for high and low proficiency students (RQ2)

A comparison between the mean number and quality of the feedback comments that High Proficiency students (HPs) and Low Proficiency students (LPs) provided and received (see Table 4) reveals that there were small expected trends (although not statistically significant) in the

Table 4

Mean number and quality of comments (and SD) provided and received by high versus low proficiency students, along with the effect size of HPs vs. LPs contrast (n = 50).

Comment type		HPs (n = 25)		LPs (n = 25)		High vs. Low	
		Mean	SD	Mean	SD	Cohen's d	p
Provided	<i>Problems</i>	13.3	11.6	14.4	10.5	0.10	ns
	<i>Suggestions/Solutions</i>	15.3	11.2	11.1	11.6	0.37	ns
	<i>Quality of feedback</i>	4.2	0.6	4.1	0.4	0.22	ns
Received	<i>Problems</i>	12.5	8.2	15.2	7.2	0.35	ns
	<i>Suggestions/solutions</i>	13.4	9.3	13.0	7.0	0.05	ns
	<i>Quality of feedback</i>	4.1	0.6	4.2	0.7	0.23	ns

overall frequency of comments provided and received: (1) HPs provided more feedback, especially more *Suggestions/Solutions* and better quality feedback (2) LPs received more feedback, especially more *Problems* and better quality feedback.

Although it is often assumed that language proficiency influences performance in providing and receiving feedback (Allen & Katayama, 2016), the patterns found in t-tests reveal that there were not large effects of L2 proficiency on the feedback features and quality provided and received. None of the high vs. low contrasts were statistically significant (all ps > 0.1 and most >0.2), and the effect sizes were all weak (d <0.5). The relatively high standard deviations for the number of comment measures suggest factors other than learner L2 proficiency (e.g., engagement with evaluation criteria, motivation for peer feedback, within-group differences in writing ability and group dynamics) might be important factors influencing amount and quality of comments (Cao et al., 2019). However, different from Cao et al. (2019), The F, p, and η² values of Author by Reviewer L2 Proficiency ANOVA interaction analysis reveal that, with both the two feedback features and feedback quality, the patterns of results across provided vs. received and across measures were very similar: weak effect sizes and no significant interaction term in all four matching groups (high×low; high×high;

low×low; low×high). In other words, L2 proficiency has no impact on provided and received feedback features and quality across all matchings. Therefore, in this context of EFL learners writing literature reviews, students of high and low L2 proficiency are equally likely to be helpful (or unhelpful) in peer review.

But what helps high and low proficiency students improve in the peer review process? Both HPs and LPs showed higher writing quality and significant improvement in the second draft and the amount of gain was roughly similar in the 8.2 % (HPs: t = 6.6***; Draft 1: M = 73.0; SD = 5.5; Draft 2: M = 79; SD = 7.9) to 10.1 % (LPs: t = 9.5***; Draft 1: M = 68.1; SD = 6.4; Draft 2: M = 75; SD = 5.9) range. In consistency with learner language proficiency difference, in both draft 1 (t = 2.7*) and draft 2 (t = 2.1*), within group difference was significant between HPs and LPs.

Comparison among the three tests (See Table 5) within the stepwise model and the full model demonstrates that, similar to the trend of the overall data, the joint predictive strength of providing and receiving (as shown in R² value) was highest for both HPs and LPs in contrast to the case when either providing or receiving was controlled. The only exception was with LPs in which R² remained the same in test 1 and test 3, but the predictors changed from providing alone (*Problems* and *Suggestions/Solutions*) to providing + receiving (*Suggestions/Solutions*).

As was the case overall, the baseline role of Draft 1 on Draft 2 quality (not shown in the table) was strong in both cases: high proficiency β = 0.69, p < .001 in the stepwise model test 3 (explaining 48 % of the variance in Draft 2) and β = 0.72, p < .001 in the full model test 3; and low proficiency β = 0.81, p < .001 in the stepwise model test 3 (explaining 66 % of the variance in Draft 2), and β = 0.85, p < .001 in the full model test 3. Therefore, the features and quality of providing and receiving feedback strongly predicted revision improvement, accounting for 87 % (i.e., (93–48 %)/(1 – 48 %)) of change in Draft 2 rating for HPs and 65 % (i.e., (88 %–66 %) / (1–66 %)) for LPs.

Looking at the particular comment predictors, the strength of providing + receiving feedback was strongest with *Suggestions/Solutions* for both HPs and LPs. In particular, providing *Suggestions/Solutions* significantly predicted revision improvement for both HPs (3 models) and LPs (2 models), and receiving *Suggestions/Solutions* predicted revision improvement for HPs (2 models) and LPs (1 model). Identifying

Table 5

The strength (partial r/standardized β) of the relationship of each providing and receiving comment predictor to Draft 2 quality, separately for HPs and LPs, controlling for Draft 1.

Predictors	Partial r with Draft 2	Regression β predicting draft 2					
		Stepwise model			Full model		
		Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
High Proficiency Students							
Provided (n = 25)							
<i>Problems</i>	0.31 [†]	-	-	-	.07	-	0.14 [†]
<i>Suggestions/Solutions</i>	0.74***	0.45***	-	0.40***	0.41**	-	0.33***
<i>Feedback quality</i>	0.34*	-	-	-	0.01	-	0.02
Received (n = 25)							
<i>Problems</i>	0.09	-	-	-	-	-0.02	0.14
<i>Suggestions/Solutions</i>	0.26	-	0.31**	0.24**	-	0.34*	0.16
<i>Feedback quality</i>	-0.28 [†]	-	-	-	-	-0.14	-0.00
R ²	NA	0.86	0.78	0.91	0.86	0.80	0.93
Max VIF	NA	1.19	1.00	1.24	1.67	2.14	2.67
Low Proficiency Students							
Provided (n = 25)							
<i>Problems</i>	0.28 [†]	.25*	-	-	0.18	-	0.17
<i>Suggestions/Solutions</i>	0.44*	0.24*	-	0.25*	0.11	-	0.12
<i>Feedback quality</i>	0.37*	-	-	-	0.22*	-	0.26*
Received (n = 25)							
<i>Problems</i>	-0.09	-	-	-	-	0.11	0.14
<i>Suggestions/Solutions</i>	-0.15	-	-	0.24*	-	0.10	0.11
<i>Feedback quality</i>	0.20	-	-	-	-	-0.01	-0.05
R ²	NA	0.83	0.69	0.83	0.84	0.71	0.88
Max VIF	NA	1.06	1.00	1.06	1.86	1.23	2.24

Notes. [†] = p < .1, * = p < .05, ** = p < .01, *** = p < .001

Problems marginally contributed to revision for HPs and LPs only in the case of providing (for HPs in partial correlation and in the full model, and for LPs in partial correlation). The quality of provided feedback was a significant predictor of Draft 2 scores for LPs in two models, and for HPs in both providing and receiving with partial correlation. In other words, the quality of provided feedback significantly predicted revision quality for LPs, whereas HPs were influenced by both the quality of feedback provided (positively) and received (negatively).

In summary, all students benefited from the overall peer review process. Although in general, providing feedback contributed more than receiving feedback and feedback features appeared to be more predictive than feedback quality, for students with different L2 proficiency, the more specific effects of providing and receiving feedback were different. HPs appeared to benefit from both providing and receiving *Suggestions/Solutions*, as well as from providing and receiving *Problems* to some extent, and they were generally not influenced by the quality of provided or received feedback. LPs were influenced by providing and receiving *Suggestions/Solutions* and by providing *Problems*. In particular, the quality of feedback provided significantly contributed to their own writing improvement.

5. Discussion

5.1. The mutual benefits of being both a feedback provider and receiver

As previously discussed, although the benefits of providing feedback (Cho & Cho, 2011; Patchan & Schunn, 2015; Yu, 2019) and receiving feedback (Gao, Schunn, & Yu, 2019; Patchan & Schunn, 2016) have been supported by many studies, the joint effect of providing and receiving feedback in the bilateral process of peer review (Wu & Schunn, 2020) largely remains unknown. This quasi-experimental study, situated in an authentic L2 learning environment, advances this line of research by exploring the learning benefits in combination of providing and receiving feedback in a complex academic writing task (writing literature reviews). Based on careful coding and statistical analysis, the study confirms and reinforces the value of peer review in L2 writing.

The joint effect of providing and receiving peer comments for all learners in this study reveals that this bi-directional peer review activity breaks the boundary of traditional one-directional instruction held by teachers through the helpful socio-interactive environment where learners get scaffolding from peers, and learners can work in a small team and undertake the responsibility for everyone's improvement through cooperative learning (Gokhale, 1995). Similar to findings by Yu and Wu (2016), being both a feedback provider and receiver are found to be mutually beneficial. As both evaluating other's work and constructing feedback (providing feedback) and processing and assessing the usefulness of feedback (receiving feedback) require reference to the rubric, engaging in the tasks associated with the role of a feedback-provider should help in the successful fulfillment of those associated with a feedback-receiver, and vice versa (Yu & Wu, 2016). This result is also consistent to students' perception of benefits in peer review in Cao et al.'s study (2019) in that the largest proportion of students believed that they benefited from both providing and receiving peer feedback. Ideally, when students' positive attitudes to peer review are further consolidated through empirical evidences in practices, there is a better chance for students to be actively engaged in this significant pedagogical activity.

5.2. High proficiency students gaining prominently from the bilateral process of providing and receiving peer feedback features/quantity

Abundant literature has established that peer review may provide benefits for students of different L2 proficiency (Allen & Mills, 2016), and that L2 proficiency significantly influences the effects of peer review (Hyland & Hyland, 2006). To further understand the joint impact of providing and receiving peer feedback on different L2 proficiency

students, this study analyzed this impact respectively for high and low proficiency students. Prior studies have noted concerns about low proficiency L2 students with respect to peer feedback (Yu & Lee, 2016b) or found differential benefits for low proficiency L2 students (Lundstrom & Baker, 2009). Interestingly, in this context of EFL academic writing with intermediate-advanced L2 proficiency students, L2 proficiency has little impact on learner performance in providing and receiving features and quality of feedback. But the gains in the process varied across student proficiency groups.

High proficiency students produced better quality drafts, but in complex tasks like writing literature reviews, they are also often challenged with high level issues of content and logic that they cannot handle on their own (Gao, Schunn, & Yu, 2019). In providing and receiving peer feedback, HPs with more sophisticated means-ends manners might embrace more competence in dealing with any problem and they can learn and revise their draft equally well using the feedback from low reviewers and high reviewers (Patchan & Schunn, 2016) and therefore made progress on their own revised drafts. Through providing *Suggestions/Solutions* and *Problems*, they practiced analytical and reasoning skills, reflected on their own writings, raised awareness of the genre of literature review writings, improved writing skills and hence worked out better text quality in revision (Cho & MacArthur, 2011; Nicol et al., 2014; Yu, 2019); while in receiving *Suggestions/Solutions*, they learned to detect and correct errors, as well as anticipate and prevent possible mistakes as they are more able to uptake the feedback received, and use it in future tasks (Ion et al., 2018). The fact that HPs were generally not influenced by the quality of feedback provided and received also indicates that providing and receiving peer feedback is rewarding to high proficiency students' own writings regardless of the good or bad quality of feedback provided and received since students with better L2 proficiency are usually the ones who are more competent and take learning (peer feedback in this case) more seriously.

5.3. Low proficiency students benefiting more from providing peer feedback feature/quantity and quality

Although L2 proficiency in peer review may be a significant factor in determining students' ability to give and utilize peer feedback, its impact on peer feedback has been underexplored (Yu & Lee, 2016b). The limited body of research seems to suggest that low proficiency students benefit more by receiving feedback and they are usually portrayed as less able to contribute to peer feedback and that they primarily play the role of a receiver rather than a giver (Allen & Mills, 2016; Lundstrom & Baker, 2009). However, low proficiency students in Yu & Lee's study (2016b) contributed to peer feedback by providing a range of comments on various aspects of writing and helped their peers enhance their writing quality.

In this study, L2 proficiency did not determine the quantity and quality of peer feedback provided or received, and low proficiency students made significant improvement in revision. It indicates that anonymity, random pairing and multiple feedback design probably encouraged all students, especially LPs (Huisman et al., 2017) as they reviewed writings of similar and/or different quality, and that learners across the class probably have ZPDs that are well within each other's reach (Lantolf & Thorne, 2006), possibly because their L2 proficiency and knowledge domain are relatively in the same range.

Generally, low proficiency students had more problems with their initial writings and more prominently so in complex writing tasks. However, different from the common hypothesis and earlier findings that LPs are less competent in providing feedback and benefit more from receiving feedback (Patchan & Schunn, 2015), LPs benefited more from providing feedback in this study. Similar findings as in Yu and Lee (2016b), Ion et al. (2018) and Yu (2019) reinforce the importance of active learning in reviewing others' work. In complex writings, improvement of document quality is more a matter of content

organization and persuasion of writing. Therefore, receiving comments alone did not lead to learning. In support of the learning by reviewing hypothesis, engaged reviewing (providing and receiving key feedback features and good quality feedback) through seeing good or bad examples, diagnosing problems, and giving meaningful suggestions or solutions (Patchan & Schunn, 2015) constituted real learning experiences and therefore are more helpful to students. The fact that providing helpful feedback to others significantly benefited LPs in revision again enforces that engaged learning in peer review is not only beneficial to feedback receivers, it creates more learning for feedback providers.

6. Conclusion

In summary, this study adopted a quantitative method to explore the joint impact of providing and receiving peer feedback on L2 students in terms of key feedback features and feedback quality in academic writing. The analyses suggest that providing and receiving feedback jointly contributed to revision more than either providing or receiving peer feedback. However, across the tested models, the predictive strength of providing and receiving feedback varied. Similar to previous findings (Cho & Cho, 2011; Berggren, 2015), providing feedback (both quantity and quality) contributed more to revision than that of receiving feedback, and some feedback features of provided and received feedback are especially beneficial whereas feedback quality played little role in gains. In particular, providing and receiving constructive feedback, that is Suggestions/Solutions, are most beneficial to all students. This finding highlights the significance of understanding more about the bilateral benefits of providing and receiving peer feedback in L2 writing context.

Pedagogically, this study further proves that peer review is an effective instructional strategy in process writing approach. In practice, teachers could perform “feedback on feedback” practice to scaffold students to include key feedback features/quantity and good quality feedback when reviewing peer’s work, and teachers should encourage students to place more emphasis on the process of providing feedback instead of the outcome (Yu, 2019). On the other hand, fostering the development of feedback literacy is a long-term endeavor. Student feedback literacy denotes the understandings, capacities and dispositions needed to make sense of information and use it to enhance work or learning strategies, specifically, in appreciating feedback, making judgments, managing affect, and taking action (Carless & Boud, 2018). In other words, capacities and dispositions in providing and receiving feedback are both necessary in peer review activities, and therefore, both should be catered to by instructors.

With regard to L2 proficiency, there were no robust main effects of author or reviewer proficiency on the amount of feedback features and the quality of comments provided or received. Therefore, by following a well-designed rubric, even in complex content-oriented writing tasks, learners of both high and low L2 proficiency can improve their assessment skills, and have an opportunity to find out problems and correct their own writings (Cho & Cho, 2011; Cho & MacArthur, 2011; Nicol et al., 2014) through providing and receiving feedback. Prominently, the bilateral benefits of providing and receiving feedback features (Suggestions/Solutions, in particular) were more significant for high proficiency students, whereas the quality of received or provided feedback did not influence their revision. However, low proficiency students were more affected by providing feedback features (Suggestions/Solutions and

Problems) and quality in writing improvement.

This finding contributes to the line of studies on how L2 proficiency can impact peer feedback (Allen & Mills, 2016; Yu & Lee, 2016b). Practically, in peer feedback training (Cui, Schunn, & Gai, 2022), consideration should be given to both features and quality of feedback while taking learner L2 proficiency and task complexity into account. In particular, in L2 contexts, low proficiency students should be encouraged to provide more constructive feedback like Suggestions/Solutions, and better quality feedback so that they would benefit more in their own revision through such engaged learning practices.

7. Limitations and future work

First, the result of this study could have been shaped by the assignment context and range restriction. The detailed writing and reviewing rubric, multiple peer feedback, and back-evaluations were designed to encourage the production of a greater quantity and better quality of feedback, but they could have influenced learner performance in general. For example, the content-oriented rubric resulted in the Problems and Suggestions/solutions being mostly on content rather than language issues. Although the ability range of the participants (L2 proficiency and domain knowledge) was large enough to show statistically significant variation in predictive effects on performance, a broader inclusion of participants from different level universities and different majors might have brought forth different results. Therefore, the result of this study should be cautiously generalized into other learning contexts and other samples.

Second, although the feedback analyzed were implementable ones, the intermediate effects of implementation on revision and factors leading to implementation for high and low L2 proficiency learners were unknown, leaving room for further investigation. Future studies should also seek for different or more comprehensive approaches to measuring feedback quality, such as using expert ratings of feedback quality.

Lastly, although the present study contributes to our understanding of the bilateral benefits of learners in peer feedback, the study was correlational in nature. To test the causality of feedback and the interplay of L2 proficiency on revision, intervention studies need to be conducted, and findings in studies like the current one can help identify focal targets for such intervention studies. A future study could, for example, experimentally manipulate whether authors received or reviewers provided certain kinds of feedback to examine the effects on authors and reviewers. If the findings replicate in experimental studies, then a number of practical recommendations follow. For example, analyzing representative models of helpful and unhelpful comments and offering a checklist of DOs and DON'Ts to guide students to give constructive as well as encouraging feedback.

Disclosure statement

The 3rd author (Christian D. Schunn) is a co-inventor of the *Peer-ceptiv* system. The other two authors have no conflict of interest.

Funding

This work was supported by National Social Science Fund of China (CN) under Grant number 18BYY114.

Appendix

Writing and Reviewing Rubric.

This task asks you to write an essay on the topic "A Critical Review of Literature on ###". In this essay, you are to show your understanding of one topic in intercultural communication through reviewing at least 10 early literature abroad and 10 at home with at least 2000 words. The review should summarize, analyze, synthesize the available research, and find out a research gap or an unanswered question in the literature. In addition, it should include a critique that assesses or weighs up the value of relevant theories, ideas, claims, research designs, methods or conclusions.

Dimension 1 Introduction

- Commenting Prompt

In this dimension, an introduction to the essay should be given. The introduction should address the topic, establish its importance and briefly describe the organization of the essay in one or two paragraphs. In addition, concerns should be given to clarity of writing, logic and organization, effectiveness/persuasiveness of writing, and English writing convention (APA style).

- Rating Criteria

Select from the following scale and give a rating according to the description.

7 - Excellent - Introducing the central topic, importance of central topic, necessity of further investigation very clearly and precisely, in a highly well-understood manner.

6 - Very good.

5 - Good - Introducing the central topic, importance of central topic, necessity of further investigation clearly in an easy-to-follow manner.

4 - Ok.

3 - Fair - Introducing the central topic, importance of central topic, necessity of further investigation.

2 - Very poor.

1 - Difficult to read at all.

Dimension 2 Body Part

- Commenting Prompt

Dimension 2 is the body part of the essay. It should analyze and synthesize the literature domestic and abroad. The following questions serve as a checklist: Did the paper review at least 20 references domestic and abroad in total? Did the paper clarify the arguments and the argument involvement in the literature? Did the paper clarify the relationships of the arguments and counter arguments? Did the paper critique the argument, research design, methodology or conclusions? Did the paper have a clear critical stance in the review? In addition, concerns should be given to clarity of writing, logic and organization, effectiveness/persuasiveness of writing, and English writing convention (APA style).

- Rating Criteria

Rate this part according to the scales below.

7 - Excellent - Present the literature in a highly critical and objective manner, with clear focus and problem involvement by comparison and contrast.

6 - Very good - Present the literature in a critical and objective manner, with clear focus and problem involvement by comparison and contrast.

5 - Good - Present the literature with clear focus and problem involvement by comparison. and contrast.

4 - Ok - Present the literature with clear focus.

3 - Fair - Present the literature by comparison and contrast.

2 - Poor - List the literature with no clear focus.

1 - Difficult to read at all.

Dimension 3 Conclusion

- Commenting Prompt

This is the conclusion part of the essay. It should conclude the essay with logical conclusions on basis of the discussion in the body part of the essay. Use the following questions as a checklist. Did the paper address all the key words and issues in the central topic? Did the paper locate and present a gap in the literature for further research? Did the paper clarify a point of departure from previous studies? Was the point of departure for further study operational and manageable for academic research? In addition, concerns should be given to clarity of writing, logic and organization, effectiveness/persuasiveness of writing, and English writing convention (APA style).

- Rating Criteria

Rate the conclusion according to the scales below.

7 - Excellent - End the literature review with an identified gap, very operational and manageable for further research.

6 - Very good.

5 - Good - End the literature with an identified gap or point of departure in a logical way.

4 - Ok.

3 - Fair - End the literature review with an identified gap.

2 - Very poor.

1 - Difficult to read at all.

References

- 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. (2016). The impact of second language proficiency in dyadic peer feedback. *Language Teaching Research*, 20(4), 498–513.
- Berggren, J. (2015). Learning from giving feedback: A study of secondary-level students. *ELT Journal*, 69(1), 58–70.
- Cao, Z., Yu, S., & Huang, J. (2019). A qualitative inquiry into undergraduates' learning from giving and receiving peer feedback in L2 writing: Insights from a case study. *Studies In Educational Evaluation*, 63, 102–112.
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325.
- Chang, C.-Y., Lee, D.-C., Tang, K.-Y., & Hwang, G.-J. (2021). Effect sizes and research directions of peer assessments: From an integrated perspective of meta-analysis and co-citation network. *Computers & Education*, 164, Article 104123.
- Cho, K., & MacArthur, C. (2011). Learning by reviewing. *Journal of Educational Psychology*, 103(1), 73–84.
- Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers and Education*, 48(3), 409–426.
- Cho, Y. H., & Cho, K. (2011). Peer reviewers learn from giving comment. *Instructional Science*, 39(5), 629–643.
- Chong, I. (2017). How students' ability levels influence the relevance and accuracy of their feedback to peers: A case study. *Assessing Writing*, 31, 13–23.
- Cote, R. A. (2014). Peer feedback in anonymous peer review in an EFL writing class in Spain. *Gist Education and Learning Research Journal*, 9, 67–87.

- Cui, Y., Schunn, C. D., & Gai, X. (2022). Peer feedback and teacher feedback: a comparative study of revision effectiveness in writing instruction for EFL learners. *Higher Education Research & Development*, 41(6), 1838-1854. <https://doi.org/10.1080/07294360.2021.1969541>
- Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review*, 32, 481-509.
- Gao, Y., Schunn, C. D., & Yu, Q. (2019). The alignment of written peer feedback with draft problems and its impact on revision in peer assessment. *Assessment & Evaluation in Higher Education*, 44(2), 294-308.
- Gokhale, A. A. (1995). Collaborative learning enhances critical thinking. *Journal of Technology Education*, 7(1), 22-30.
- Huisman, B., Saab, N., Driel, J. V., & 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(7), 1433-1447.
- Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83-101.
- Ion, G., Sánchez Martí, A., & Agud Morell, I. (2018). Giving or receiving feedback: Which is more beneficial to students' learning. *Assessment & Evaluation in Higher Education*, 39(1), 102-122.
- Kwan, B. S. C. (2006). The schematic structure of literature reviews in Doctoral theses of applied linguistics. *English for Specific Purposes*, 25(1), 30-55.
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.
- Li, L., Liu, X., & Steckelberg, A. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology*, 41(3), 525-536.
- Li, L., Liu, X., & Zhou, Y. (2012). Give and take: A re-analysis of assessor and assessee's roles in technology-facilitated peer assessment. *British Journal of Educational Technology*, 43(3), 376-384.
- Liu, J., & Hansen, E. J. (2002). *Peer response in second language writing classrooms*. Arbor: University of Michigan Press.
- Liu, N. F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279-290.
- 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. (2008). Reviewer stances and writer perceptions in EFL peer review training. *English for Specific Purposes*, 27(3), 285-305.
- Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, 37(4), 375-401.
- Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102-122.
- Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feedback: How students respond to peers' texts of varying quality. *Instructional Science*, 43(5), 591-614.
- 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), 227-265.
- Patchan, M. M., Schunn, C. D., & Clark, R. J. (2018). Accountability in peer assessment: Examining the effects of reviewing grades on peer ratings and peer feedback. *Studies in Higher Education*, 43(12), 2263-2278.
- Slavin, R. E. (1990). *Cooperative learning: Theory, research and practice*. Englewood Cliff, NJ: Prentice Hall.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge, MA: Cambridge University Press.
- Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge, MA: Cambridge University Press.
- Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276.
- Topping, K. J. (2009). Peer assessment. *Theory into Practice*, 48(1), 20-27.
- Trautmann, N. M. (2006). *Is it better to give or to receive?: Insights into collaborative learning through web-mediated peer review*. (PhD diss.). Cornell University.
- Winstone, N. E., Boud, D., Dawson, P., & Heron, M. (2022). From feedback-as-information to feedback-as-process: a linguistic analysis of the feedback literature. *Assessment & Evaluation in Higher Education*, 47(2), 213-230.
- Wu, Y., & Schunn, C. D. (2020). From feedback to revisions: Effects of feedback features and perceptions. *Contemporary Educational Psychology*, 60, 101826.
- Wu, Z. (2019). Lower English proficiency means poorer feedback performance? A mixed-methods study. *Assessing Writing*, 41, 14-24.
- Yu, F. Y., & Wu, C. P. (2016). Predictive effects of the quality of online peer-feedback provided and received on primary school students' quality of question-generation. *Educational Technology & Society*, 19(3), 234-246.
- Yu, S. (2019). Learning from giving peer feedback on postgraduate theses: Voices from Master's students in the Macau EFL context. *Assessing Writing*, 40, 42-52.
- 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.
- Zhang, X. (2018). *An examination of the effectiveness of peer feedback on Chinese university students' English writing performance*. (Ph.D. Dissertation of Oakland University). Rochester: Oakland University.
- Zong, Z., Schunn, C. D., & Wang, Y. (2020). Learning to improve the quality peer feedback through experience with peer feedback. *Assessment & Evaluation in Higher Education*, 46(6), 973-992.