



The Effects of Online Feedback on ESL/EFL Writing: A Meta-Analysis

Xiaoxuan Lv¹ · Wei Ren¹ · Yue Xie¹

Accepted: 17 May 2021 / Published online: 27 May 2021
© De La Salle University 2021

Abstract Online feedback is frequently implemented during second/foreign language (SL/FL) writing tasks and assessments. This meta-analysis investigates the effectiveness of online feedback in SL/FL writing. After careful screening and the application of inclusion and exclusion criteria, this study synthesizes the results of 17 primary studies reporting on students' English SL/FL writing quality after online feedback. The studies involved 1568 students, and the results indicate a Hedges' g effect size of 0.753 for the effectiveness of written feedback in general. Online feedback from teachers/instructors produces a larger effect size ($g = 2.248$) than online peer feedback ($g = 0.777$) and online automated feedback ($g = 0.696$). It was also found that educational levels and task genre mitigate the impact of online feedback on writing quality. Overall, the findings contribute to a better understanding of the impact of online feedback on ESL/EFL writing and provide insights into online ESL/EFL writing instruction.

Keywords Online feedback · ESL/EFL writing · Meta-analysis · Moderator analyses

Introduction

Feedback is frequently implemented in writing tasks in English as a second/foreign language (ESL/EFL) writing courses across the world. It remains a core feature of the writing classroom, since constructive feedback can raise

students' awareness, improve their texts, and help them learn to use the language effectively (Hyland & Hyland, 2019). Previous research on written feedback has focused more on face-to-face classroom instruction than on online instruction. However, in recent years, online writing classes have started to play an important role in improving students' writing skills and developing a sense of autonomy (Kourbani, 2017). Since the outbreak of COVID-19 and its spread across the world, instructional activities in universities have been severely disrupted and online teaching has become a substitute for traditional face-to-face classroom teaching. It is therefore timely to evaluate the effectiveness of online feedback on students' ESL/EFL writing. Such studies will provide useful information for teachers to devise or improve their feedback techniques, which can contribute to enhancing the learning potential of written feedback in online teaching contexts.

Online feedback in teaching contexts can be defined as the post-response information provided through online means, "which informs the learners on their actual states of learning and/or performance" (Narciss, 2008, p. 292). Accordingly, online feedback on ESL/EFL writing refers to the judgment of text provided by teachers, students, or automated software through online means, which reflects concern for learners' writing (Hyland & Hyland, 2019). Some empirical studies (e.g., Guasch et al., 2013; Latifi et al., 2019; Link et al., 2020; Noroozi & Hatami, 2019) have investigated the impact of online feedback on ESL/EFL writing, which generally reflect that online feedback contributes to the development of ESL/EFL writing and provides better pedagogical meaning for ESL/EFL teachers or learners. However, a quantitative synthesis with an accurate and holistic view across studies is needed to further examine the impact of online feedback from different sources, such as teachers, peers, and automated feedback,

✉ Wei Ren
weiren@buaa.edu.cn

¹ School of Foreign Languages, Beihang University, No. 37 Xueyuan Road, Beijing 100191, People's Republic of China

and to investigate how this feedback improves students' writing quality as an essential part of teaching activities. Writing quality in this paper is operationalised as the performance of learners' writing product in terms of complexity, accuracy, fluency, content, organization, and proper use of linguistic and textual parts of the language, etc. (Cumming et al., 2000; Housen & Kuiken, 2009; Sasaki, 2000). Therefore, a comprehensive research synthesis will be conducted with the meta-analysis, one of the most effective tools for research synthesis (Li, 2010), to explore the effectiveness of feedback in online ESL/EFL writing courses on fostering learners' writing quality.

Literature Review

Meta-Analyses on Written Feedback

Meta-analyses have become increasingly popular in the field of second language (L2) writing, since these quantitative approaches to averaging effect sizes across studies can be systematic and replicable, with additional strengths such as increased statistical power, moderator analyses, etc. (Oswald & Plonsky, 2010). A number of meta-analyses about the effect of feedback on L2 writing have been conducted, particularly related to corrective feedback (Biber et al., 2011; Chen & Renandya, 2020; Huisman et al., 2019; Kang & Han, 2015). For example, Biber et al. (2011) investigated the effects of various types of feedback on the quality of students' writing, examining 23 studies involving both written and oral feedback. They found that the effect of written feedback in studies utilizing a pre-/post-test design ($g = 0.68$) was larger than in those employing a treatment/control design ($g = 0.40$), and there was little overall difference between teacher feedback and feedback from other sources (peer, self, and computer combined).

Synthesizing 21 primary articles, Kang and Han (2015) found that learners in a second language setting gain more from written feedback than those in a foreign language setting. Narrative and descriptive writing tasks showed significant differences between separate groups within this genre, so they suggested that L2 instructors should be aware that not all types of writing are as easy to correct as narrative and descriptive texts. In a more recent study Chen and Renandya (2020) investigated 35 primary studies and found an overall effect size of $g = 0.59$, indicating a positive influence of written corrective feedback on L2 written grammatical accuracy. It was found in this analysis that learners' proficiency was the strongest mitigator.

Different from the above studies, Huisman et al. (2019) synthesized 24 quantitative studies reporting on higher education students' academic writing. It was found that

engagement in peer feedback led to greater writing improvements compared to controls and self-assessment. There was no significant difference in students' writing improvement after peer feedback or teacher feedback. In addition, the authors argued that there were few well-controlled studies in this field, and stated that more methodologically sound research is needed.

The findings of previous meta-analyses indicate the effectiveness of corrective feedback or peer feedback on L2 writing. However, these studies center on written feedback provided by teachers or peers in face-to-face classroom teaching, giving little attention to online feedback and feedback source as a potential moderating variable. With the assistance of technology and the Internet, online feedback on language writing has many advantages in terms of information storage, multimodality, user friendliness, instant access to feedback, and increased interaction compared to traditional teacher and student feedback in classrooms (Bakla, 2020; Chen, 2016). As online learning has become increasingly popular and important for language education, particularly amid the disruption caused by the COVID-19 pandemic, it is necessary to explore whether online feedback can effectively improve ESL/EFL writing quality.

Online Teacher Feedback

Given the relevance of electronic-mediated instructions, writing teachers, particularly those that work in college settings, are providing more online feedback via electronic files, chats, wikis, and blogs (Elola & Oskoz, 2017; Hyland & Hyland, 2019). Empirical evidence suggests that students can benefit from electronic writing feedback provided by instructors, since there is a significant difference between the quantity and quality of conventional and electronic feedback (Johnson et al., 2019).

Research has also examined student attitudes toward online teacher feedback, suggesting that students prefer to obtain online feedback due to its convenience and high quality compared to handwritten feedback (McCabe et al., 2011; McGrath & Atkinson-Leadbetter, 2016). Additionally, it was found that online teacher feedback was more likely to be content-focused, although attention was also paid to grammar and language use (Ene & Upton, 2018).

Online Peer Feedback

Empirical studies of online peer feedback demonstrate that peer feedback can provide better local and global revisions (Yang, 2016; Yang & Meng, 2013), and improve students' writing quality (Huisman et al., 2019; Noroozi & Hatami, 2019; Noroozi et al., 2016; Novakovich, 2016; Pham et al., 2020). In addition, some studies suggest that peer feedback

can help students to grasp domain-specific knowledge (Latifi et al., 2019; Noroozi et al., 2016). For instance, Ciftci and Kocoglu (2012) found that online peer feedback is an effective way for responding to students' writing. Teachers can gain useful information from peer feedback, which can help them to design activities that enable students to read, respond to, and revise their peers' essays.

The exploration of the relationship between online feedback and its impact on EFL writers' revisions reflects that online peer feedback has the greatest impact on students' revisions (Liou & Peng, 2009; Tuzi, 2004). This is partly due to the fact that online peer feedback is more friendly and supportive, and less face-threatening (Ma, 2019). Compared to affective feedback and metacognitive feedback in online peer review, cognitive feedback was found to be more beneficial to students' writing learning gains (Cheng et al., 2015).

However, there are also some contradictory findings. Chen (2016) qualitative synthesis found that online peer feedback may not always produce positive results. Some research investigating the quality of online peer feedback has noted that peer feedback was generally related to lower-order concerns due to students' linguistic limitations (Schultz, 2000; Tolosa et al., 2013). Other studies found that peers lacked confidence and strategies when using online feedback (DiGiovanni & Nagaswami, 2001). However, Pham (2019) found that the effects of lecturer e-comments and peer e-comments on student revisions showed no statistical differences. Comparing the effectiveness of automated corrective feedback and online peer feedback, Shang (2019) found that the latter was potentially more useful in improving sentence writing, reducing grammatical errors, and producing more types of lexical items. Vaezi and Abbaspour (2015) found there was no statistically significant difference between the effects of face-to-face and online peer written corrective feedback in terms of writing achievement.

Online Automated Feedback

Alongside the development of artificial intelligence technologies, online automated feedback (OAF) generated by writing evaluation systems is being widely adopted in ESL/EFL writing instruction. By providing instant feedback, OAF is believed to be effective in improving students' writing (Kellogg et al., 2010). Students who receive more automated feedback are likely to interact with the tool more frequently, make more writing revisions, and build bridges between their prior and desired knowledge successfully (Link et al., 2020; Morch et al., 2017; Saricaoglu, 2019). There is also evidence showing an increase in writing scores as students make diligent revisions based on automatic feedback on their essays (Zhu et al., 2020).

Cheng (2017) investigated the impact of OAF on students' reflective journals in a 13-week EFL course at the university level. He found that OAF could provide immediate feedback on strengths and weaknesses in the students' reflective writing, thereby increasing their awareness during L2 learning. Stevenson and Phakiti (2019) noted that students are likely to revise their drafts immediately after receiving OAF due to its convenience, but their engagement in the perspectives of cognition and affection was not enough to help them develop learning and writing. OAF can also help teachers to adjust the focus of their feedback, enabling them to provide more instruction time and reducing their workload (Zhang & Hyland, 2018).

As the above reviews indicate, previous research on the effects of online feedback on ESL/EFL writing has produced divergent findings. Therefore, we conducted this meta-analysis to synthesize the existing findings with regard to the effects of online feedback on writing quality. The studies included in this meta-analysis provided the overall writing score, while a few of them provided specific scores for particular criteria. No matter what approaches the included studies chosen to report the writing scores, they all assessed the writing quality for ESL/EFL learners. In addition, we also explored factors related to online feedback that may influence its effectiveness in terms of writing proficiency improvement. It is hoped that this meta-analysis will provide useful insights into online feedback in ESL/EFL writing instruction. The specific research questions are as follows:

1. To what extent does online feedback influence ESL/EFL students' writing in general?
2. Which factors may affect the effectiveness of online feedback on ESL/EFL students' writing?

Method

Inclusion Criteria and Literature Search Strategies

The following criteria were used to determine whether published studies qualified for inclusion in this meta-analysis:

1. The study employs systematic quantitative data suitable for a meta-analysis and was published in the twenty-first century, specifically between January 2000 and February 2021 (since online learning was not popular last century).
2. The study includes the instructional effects of online/automated/electronic feedback on any type of ESL/EFL writing.

3. The independent variables involve some type of reasonably well-described online feedback related to ESL/EFL writing.
4. The target language of instruction is either a second or foreign language for the study participants.

Studies were excluded from the analysis if any of the following criteria were met:

1. The study employed quantitative data, but did not report any descriptive statistics.
2. The study did not focus on writing quality, instead examining the attitudes or perspectives of students.
3. The articles were published in languages other than English.

The systematic search was conducted via SCOPUS and Web of Science. The following search items were used to interrogate the databases: “online writing, online essay, online composition” in the subject field, combined with “peer feedback, teacher feedback, instructor feedback, peer review, peer comment, teacher review, teacher comment, automated feedback, e-feedback, electronic feedback”. Totals of 39 articles from SCOPUS and 149 articles from Web of Science were retrieved. After merging duplicate articles, a final total of 177 articles was retrieved. With reference to the above criteria, screening was conducted by inspecting the title and abstract of each item; in this screening, 123 articles were identified as not relevant to the research subject. Next, a closer examination of the methods and results sections excluded a further 37 studies because of insufficient statistics for the calculation of effect sizes, and/or study objects of other non-English languages. The final result was a total of 17 quantitative studies meeting the above selection criteria.

Coding of Studies

The coding scheme for extracting study characteristics was based on common variables considered in previous meta-analysis research in applied linguistics (Norris & Ortega, 2006), and also guided by the suggestions of Lipsey and Wilson (2001). Study characteristics were coded to reflect potential moderating variables for the effects of online feedback on writing. They included sample characteristics (educational levels, major, research setting), research method variables (feedback source, study design, task type, task setting), and effect size (total sample size, treatment/control-group size/mean/standard deviation, pre-/post-test differences of means/t-values, etc.).

To carry out a reliability check (Lipsey & Wilson, 2001), the first and second authors independently coded a random 20% of the primary studies (N = 17) using a coding manual involving information on effect size

calculations and the study characteristics. Discrepancies between the two coders were resolved through discussion to add credibility to the coding process. In the second round, the researchers coded all the remaining studies and an overall agreement ratio of 0.95 was observed. Any remaining discrepancies were discussed until agreement was reached.

Selection of Effect Model

The fixed-effect model and the random-effects model are the two statistical models used in most meta-analyses. The selection of effects must be based on which model best fits the distribution of effect sizes. Since this research intends to estimate the mean of a distribution of effects rather one true effect, the random-effects model is more suitable (Borenstein et al., 2009).

Computation and Interpretation of Effect Sizes

The present study applied Comprehensive Meta-Analysis (CMA version 3.3.070) to conduct all the analyses due to its flexibility in working with many different kinds of data (Borenstein et al., 2009). Effect sizes for the study were calculated using the outcome measures of writing quality under the random-effects model. Some measures were given as continuous variables and others as correlations between writing quality and online feedback. The data format in this study involved treatment- vs. control-group designs and standardized mean change differences (pre-/post with both treatment and control groups). The means, standard deviations, and sample sizes of the treatment and control groups in each study and the correlations between pre and post-tests were teased out and used for the computation of the effect sizes.

As a result of this analysis, six studies (Ciftci & Kocoglu, 2012; Ge, 2011; Link et al., 2020; Noroozi & Hatami, 2019; Noroozi et al., 2016; Pham, 2019) yielded only one effect size. The other studies included multiple feedback sources and multiple outcome measures studies, so their effect sizes were averaged for each feedback source. In addition, two studies (Ma, 2019; Shang, 2019) yielded two separate effect sizes. Therefore, 19 effect sizes were produced from our sample of 17 primary studies. Hedges' g (a conservative version of Cohen's d) was calculated, due to its correction for biases caused by small sample sizes (Lipsey & Wilson, 2001). The benchmark for the interpretation of Hedges' g is similar to that for Cohen's d , with $d = 0.20$ considered a small effect, $d = 0.50$ a medium effect, and $d = 0.80$ a large effect (Cohen, 1988).

Eight variables were identified as moderator variables for the present study. Choices of the moderator variables

were based on recommendations from previous studies (Chen & Renandya, 2020; Huisman et al., 2019). As shown in Table 1, these variables include study setting, study design, feedback source, educational levels, writing task genre, type and setting, as well as participants' major.

Results

Overall Analysis

To address the overall effectiveness of online feedback, Hedges' *g* for the effects of online feedback on ESL/EFL writing were investigated under the random-effects model. Figure 1 shows the descriptive statistics for the 19 results from the 17 studies, including forest plots, variance, and standard errors.

In the forest plot, reviewers can identify the precision of each study by the length of the confidence interval. The overall effect size is calculated for the 19 results based on the pre-post design and treatment- vs, control-group design,

with a result of 0.753 (Fig. 1), which is medium to large (Cohen, 1988). All studies reported positive effects for online feedback. Six effect sizes in the primary studies exceeded 1.0, representing a large effect.

Publication Bias

A funnel plot is a type of mechanism for displaying the relationship between study size and effect, which can interpret potential evidence of publication bias (Borenstein et al., 2009). If publication bias is not present, the studies will be distributed symmetrically about the mean effect size due to random sampling error (Borenstein et al., 2009). Figure 2 shows that the effect sizes are distributed symmetrically around the average effect size, indicating the absence of publication bias.

In addition, Rosenthal's fail-safe *N* was calculated. The result was 1759, which indicate that 1759 studies would be needed to invalidate a significant effect size result. The number of 1759 was more than the criterion number of $5k + 10 = 105$, where $k = 19$ studies (Rosenthal, 1991).

Table 1 Characteristics of primary studies included in the analysis

Author	Sample Size	Effect Size(<i>g</i>)	Research setting	Educational levels	Major	Study design	Feedback source	Task genre	Task type	Task setting
Link et al. (2020)	12	1.609	EFL	U	EM	P	OAF	A	I	CA
Pham (2020)	40	1.986	EFL	U	OM	P	OPF	A	I	AT
Ma (2020) (1)	37	0.983	ESL	U	OM	EC	OPF	A	C	AT
Ma 2020 (2)	37	2.248	ESL	U	EM	EC	OTF	A	C	AT
Pham (2019)	57	2.021	ESL	U	EM	EC	OPF	V	I	CA
Shang (2019) (1)	129	0.253	EFL	U	EM	EC	OPF	N	I	CA
Shang (2019) (2)	129	0.523	EFL	U	EM	EC	OAF	N	I	CA
Saricaoglu (2019)	31	0.252	ESL	U	OM	P	OAF	A	I	CA
Noroozi and Hatam (2019)	42	0.483	ESL	U	OM	P	OPF	A	I	CA
Cheng (2017)	129	0.196	EFL	U	OM	EC	OAF	RJW	I	CA
Morch et al. (2017)	48	1.68	EFL	S	OM	P	OAF	N	I	CA
Noroozi (2016)	189	0.317	ESL	U	OM	P	OPF	A	I	CA
Yang (2016)	39	1.924	EFL	U	EM	EC	OPF	S	I	CA
Tai et al. (2015)	107	0.154	EFL	UP	OM	EC	OPF	V	I	CA
Cheng et al. (2015)	235	0.649	EFL	U	OM	P	OPF	R	I	CA
Vaezi and Abbaspour (2015)	40	0.801	EFL	L	OM	P	OPF	A	I	AT
Guasch et al. (2013)	201	0.873	EFL	U	OM	EC	OPF	E	I	CA
Ciftci and Kocoglu (2012)	30	0.841	EFL	U	EM	EC	OPF	A	I	CA
Ge (2011)	36	0.209	EFL	U	OM	P	OPF	E	I	AT

In the column for educational levels, *U* = university, *UP* = upper secondary school, *L* = Language Institute. For study design, *P* = pre-/post-test design, *EC* = treatment- vs control-group design. In the major column, *EM* = English major, *OM* = other majors. For feedback source, *OPF* = online peer feedback, *OAF* = online automated feedback, *OTF* = online teacher feedback. For task genre, *A* = argumentative essay, *N* = narrative essay, *R* = report, *RJW* = reflective journal writing, *S* = summary writing, *V* = various. For task type, *I* = individual task, *C* = collaborative task. For task setting, *CA* = classroom assignment, *AT* = assessment task

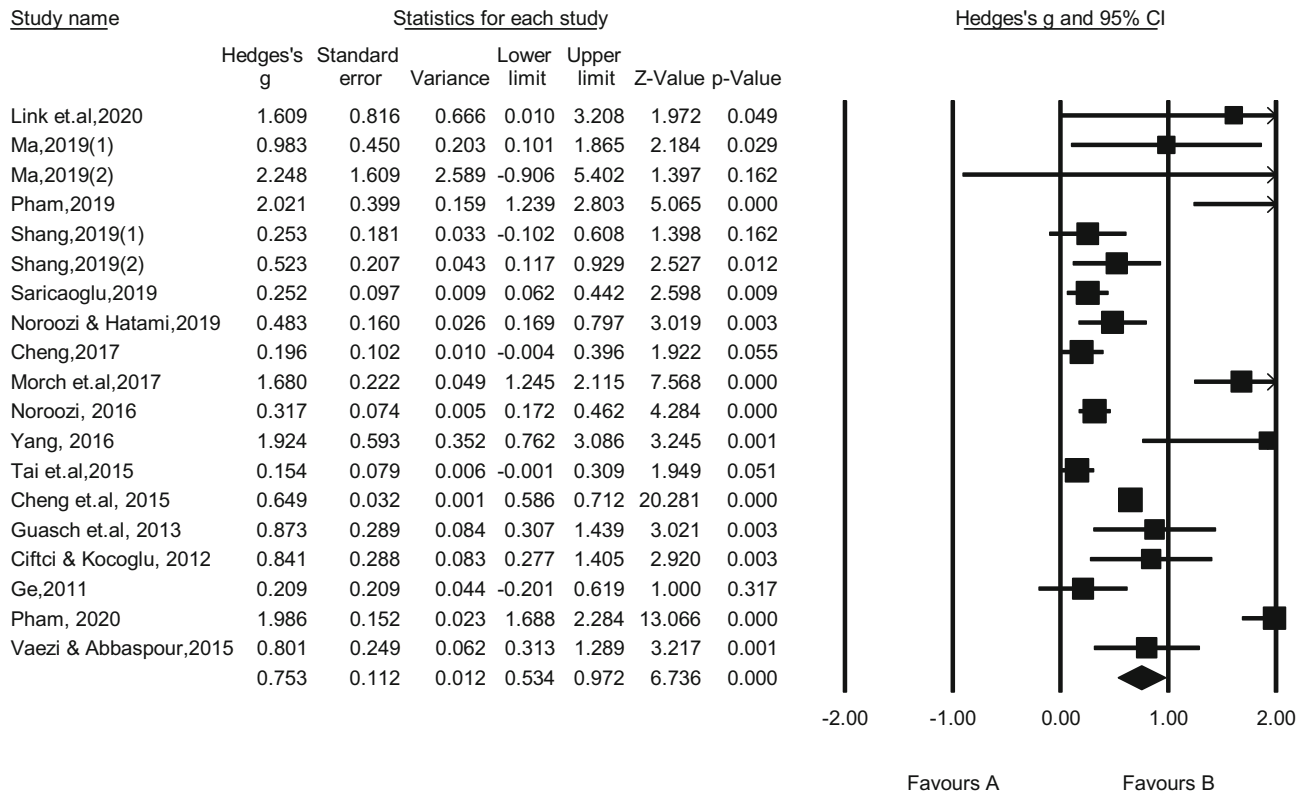


Fig. 1 Forest Plot of Effect Sizes for Overall Online Feedback

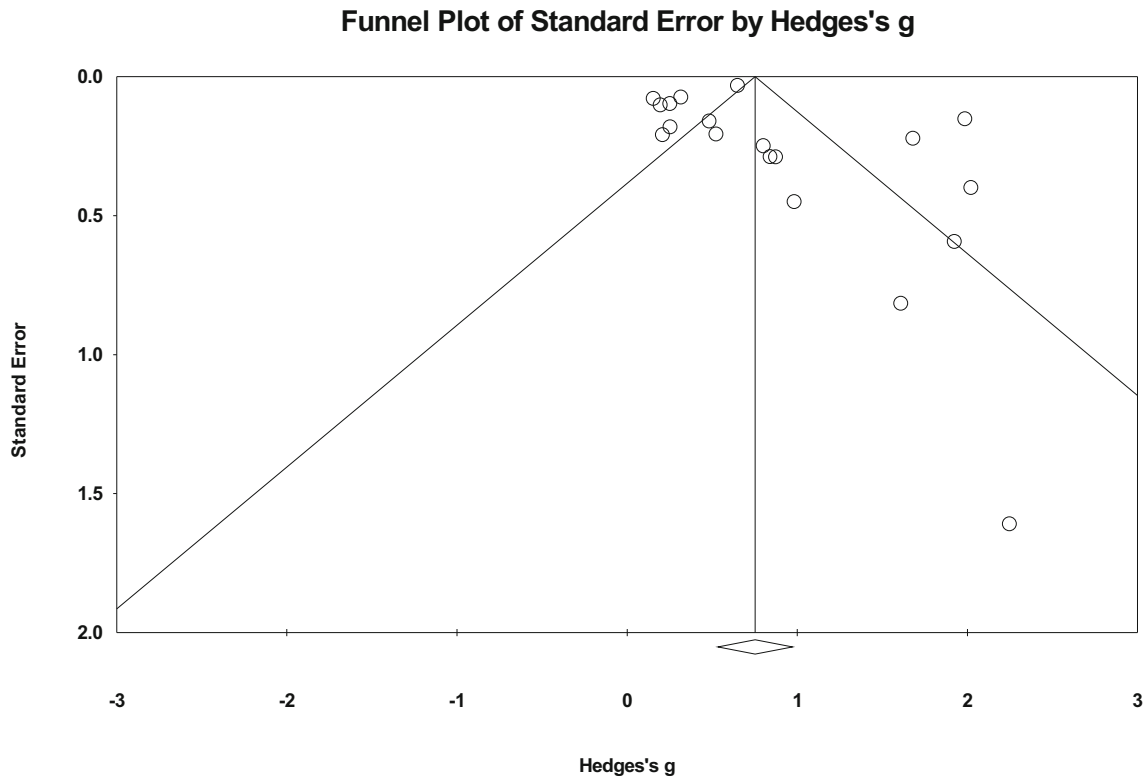


Fig. 2 Funnel Plot of Standard Error by Hedges' g

Therefore, the fail-safe N results additionally show that there was no serious publication bias, and this aspect had minimal impact on the results of the meta-analysis.

Homogeneity of Effect Sizes

The Q test for homogeneity of effect size was conducted based on the random-effects model used in the meta-analysis, and the result was $Q(18) = 212.571, p < 0.01$, indicating that the null hypothesis should be rejected. It was found that the effect sizes varied significantly across studies. The I^2 statistic was 91.532, indicating that a high proportion of the between-effect size variance reflected real differences in effect sizes (Borenstein et al., 2009). Such a result may be caused not only by sampling error but also by moderator factors (Rothstein et al., 2006).

Moderator Analysis

Moderator analysis examines the factors influencing the effects of online feedback. Borenstein et al. (2009) point out that a random-effects model cannot estimate between studies variance with precision if the sample size is small. Instead, a fixed-effect model should be applied when the sample size of the subgroup is less than five. A summary of eight moderator variables and their respective effect sizes is presented in Table 2.

In terms of study characteristics, the three variables of study setting, educational levels and major were analyzed. Online feedback showed a similar effect when provided in an ESL setting ($g = 0.796$) and an EFL setting ($g = 0.601$), but the differences were not significant ($p = 0.376$), indicating that study setting did not mediate the effect of online feedback. In terms of educational levels, studies in universities yielded a medium effect size ($g = 0.347$). Only one study was conducted in a junior college, exhibiting a small effect size ($g = 0.154$), while two studies in secondary schools yielded a large effect size ($g = 1.248$). The differences among the three levels were significant under the fixed-effect model ($p = 0.000$). Finally, there were no significant differences ($p = 0.216$) between the participants majoring in English or other subjects. However, English majors ($g = 1.061$) generated a larger effect size than other majors ($g = 0.678$).

Concerning research method and design, feedback source and writing task characteristics were analyzed separately. The type of study design did not produce any significant differences between categories ($p = 0.568$). The effect size for pre-/post-test designs ($g = 0.805$) slightly exceeded that for treatment- vs. control-group designs ($g = 0.678$). Concerning feedback source, no significant differences were observed among the three sources of feedback ($p = 0.627$). OTF had the greatest impact on

writing quality in only one study ($g = 2.248$). OAF ($g = 0.696$) and OPF ($g = 0.777$) both yielded medium effects. Regarding writing tasks, the three factors of task genre, task type and task setting were assessed. Task genres produced significant differences ($p = 0.000$), with summary writing gaining most from online feedback ($g = 1.924$), followed by various genres ($g = 1.047$) and narrative essays ($g = 0.812$). Argumentative essays ($g = 0.790$) and reports ($g = 0.649$) showed medium effects. However, reflective journals seemed to gain little from online feedback ($g = 0.196$). Writing task types did not show any significant differences ($p = 0.452$). However, the effect size for collaborative tasks ($g = 1.075$) was much higher than for individual tasks ($g = 0.760$). Similarly, analysis of the task setting did not produce any significant differences between the assessment task and classroom task ($p = 0.339$), although the effect size for the former ($g = 1.081$) was higher than that for the latter ($g = 0.625$).

Discussion

To address the first research question on the overall effect of online feedback, the study yielded an overall effect size of $g = 0.753$. With reference to Cohen's (1988) benchmark, this result is above the threshold for a medium effect size ($g = 0.5$). However, if interpreted according to the new benchmark for L2 research effect size proposed by Oswald and Plonsky (2010), this result falls within the range of small to medium ($g = 0.4$ is small, 0.7 is medium and 1.0 is large). This finding demonstrates the effectiveness of online feedback in terms of improving ESL/EFL writing quality. It supports most of the findings from previous meta-analyses (Biber et al., 2011; Kang & Han, 2015), thus improving our confidence in the effectiveness of online corrective feedback and online peer feedback.

On the other hand, the overall effect in this meta-analysis is slightly different from that found in previous analyses. For instance, analyzing peer feedback on academic writing, Huisman et al. (2019) found an effect size of $g = 0.9$. Chen and Renandya (2020) investigated the efficacy of written corrective feedback in L2 writing instruction, producing a mean effect size of $g = 0.59$. The variation between the present study and previous ones may result from feedback modes, feedback sources, or study settings. The feedback mode in this meta-analysis is online feedback, while feedback in previous meta-analyses was offered in face-to-face classroom interactions. In terms of feedback sources, this study incorporated forms of feedback as diverse as online peer feedback, online teacher/instructor feedback, and automated feedback. However, most previous studies focused on corrective feedback from

Table 2 Summary of Moderator Variables

No	Moderator Variable	<i>K</i>	Effect Size (<i>g</i>)	Lower 95% CI	Upper 95% CI	Total between <i>P</i> -value
1	Research Setting					0.376
	EFL	13	0.796	0.504	1.088	
	ESL	6	0.601	0.282	0.919	
2	Education level					0.000
	University levels	17	0.683	0.459	0.907	
	Upper Secondary	1	1.68	1.245	2.155	
	Language Institute	1	0.801	0.313	1.289	
3	Major					0.216
	English major	7	1.061	0.509	1.612	
	Other majors	12	0.678	0.425	0.931	
4	Study design		0.568			
	Treatment vs. control-group	10	0.678	0.377	0.979	
	Pre-/post-test design	9	0.805	0.487	1.124	
5	Feedback source					0.627
	OAF	5	0.696	0.211	1.182	
	OPF	13	0.777	0.513	1.041	
	OTF	1	2.248	− 0.906	5.402	
6	Task genre					0.000
	Argumentative essay	11	0.79	0.398	1.183	
	Narrative essay	3	0.812	− 0.02	1.643	
	Report	1	0.649	0.586	0.712	
	Reflective journal	1	0.196	− 0.004	0.396	
	Summary writing	1	1.924	0.762	3.086	
	Various	2	1.047	− 0.781	2.874	
7	Task type					0.452
	Collaborative Task	2	1.075	0.225	1.924	
	Individual Task	17	0.76	0.514	0.962	
8	Task setting					0.339
	Assessment task	5	1.081	0.169	1.993	
	Classroom task	14	0.625	0.424	0.826	

k = number of studies, *CI* = confidence interval

teachers; only Huisman et al. (2019) analyzed peer feedback.

With regard to the second research question on the factors affecting the effectiveness of online feedback, the study found that among the eight moderators, educational level and task genre produced significant differences. Other moderator variables did not yield significant differences, a finding that might be caused by a lack of adequate studies for these sub-groups. However, those variables did still reveal positive and large effects. For instance, students following an English major showed a larger effect of online feedback.

Although the reasons for this variance are beyond the scope of this meta-analysis, the findings may provide some implications for improving students' writing. In terms of research settings, the results seem to suggest that learners

in EFL contexts benefit from online feedback to a similar extent as those in ESL contexts. The findings also indicate that online writing feedback can generate impact regardless of the context. Regarding study design, the results show that pre-/post-test comparisons generated larger gains than studies of treatment- vs. control groups. This is consistent with a previous finding on written feedback reported by Biber et al. (2011), who considered these differences to be a result of the natural development in writing proficiency that comes with time.

The findings showed that different educational levels as a mitigator might affect writing quality after learners have received online feedback. The result showed larger effect sizes of online feedback at upper secondary school than at university and in language institutes. This may be caused by a lack of adequate studies in the separate groups; only

one study was conducted in upper secondary schools and language institutes. Thus, it is hard to generalize this finding.

This meta-analysis found that task genres such as summary writing, argumentative essays and narrative essays significantly mitigated the impact of online feedback on writing quality. This result is consistent with the findings of Kang and Han (2015). In addition, this research also shows that narrative, argumentative texts and summary texts are easier to correct after feedback. However, reflective journals yielded a small effect size, which might be related to the private characteristics of journals, meaning that suggestions for error correction are often ignored or rejected by L2 learners (Kang & Han, 2015). Consequently, EFL/ESL instructors should be aware of genre selection and combination when designing writing tasks.

However, due to concerns about the usefulness and correctness of their feedback, students often consider giving genre-based peer feedback to be difficult and challenging (Yu, 2020). Thus, training for providing genre-based feedback is key to improve the effectiveness of peer feedback and its impact on writing quality. Additionally, in terms of task setting, the tasks in the assessment can have a significant high effect. Therefore, ESL/EFL instructors should select or combine the above genres in their assessment tasks in order to improve students' writing.

Regarding feedback sources, the results showed that OTF, OPF and OAF all yielded large effect sizes. It seems that OPF can generate slightly more gains than OAF; this result is consistent with Shang's (2019) findings indicating that the feedback that occurred in OPF was more usable than that in OAF, which may prompt learners to write more sentences, make fewer grammatical errors, and produce more lexical items and types of words. Since studies on teacher online feedback on L2 writing are quite scarce (Ene & Upton, 2018), there was only one effect size for OTF, but it yielded the highest effect size. This result might be due to the higher reliability of teacher feedback in comparison with peer feedback as perceived by students, which may facilitate them in revising and improving their writing (Ertmer et al., 2007). Tai et al. (2015) and Yang et al. (2006) also noted that students were concerned about the quality of peer feedback. To solve these problems and improve students' confidence in peer feedback, effective feedback training should be provided by instructors, which in turn would improve EFL college students' text revisions (Berg, 1999; Min, 2005, 2006; Yang et al., 2006).

Collaborative writing refers to an interactive activity that two or more learners co-construct knowledge and produce one text (Elola & Oskoz, 2010; Storch, 2013). The analysis indicates that collaborative writing tasks can reap greater gains from online feedback than individual tasks. Previous findings indicated that collaborative writing

significantly improved the overall writing performance of EFL learners compared to individual writing (Shehadeh, 2011). In other words, students may produce better texts in collaborative writing tasks than in individual tasks (Pae, 2011). The result of task type analysis is consistent with the empirical finding reported by Guasch et al. (2013) that online feedback affects students' writing performance in collaborative writing positively, which may be related to effective interactions between learners during collaborative writing. In addition, Ma (2019) suggested that critical comments, particularly in OPF, are key to the quality of the final collaborative writing. Therefore, collaborative writing as a learning pedagogy (Onrubia & Engel, 2009) may be a better choice for ESL/EFL instructors to improve the learners' writing quality by interacting with peer feedback in online teaching environments.

Conclusions

This meta-analysis has synthesized the impact of online feedback on L2 writing. The overall effect size was small to medium, suggesting a positive impact of online feedback. The moderator variables of feedback source, task genre, and teaching status were the most effective mitigating factors. The findings indicate that the provision of online feedback training should be given serious consideration and offer implications for the instruction of online L2 writing, especially in terms of classroom task design.

This study has some limitations. First, the sample size was quite small, and the sample sizes of the sub-groups were not equal. Second, a couple of categories in some moderator variables (e.g., task genres, online teacher feedback) only drew data from a single study. Therefore, the results should be interpreted carefully. Third, some variables such as students' age, English proficiency, and outcome measures were not analyzed because the relevant information was not available in primary studies, or the variables were present with too many sub-groups in inadequate studies.

In light of these limitations, the meta-analysis illustrates some areas in need of further investigation. First, it is recommended that future research on online feedback in L2 writing includes clear descriptions of students' L2 proficiency levels and biographical information for the participating students. Additionally, future research could combine meta-analysis with a systematic review of online feedback features, which would lead to more comprehensive and insightful conclusions.

Funding The study was partially supported by the National Social Science Fund of China (20BYY066).

Data Availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code Availability Software application: Comprehensive Meta-Analysis 3.3.070.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Bakla, A. (2020). A mixed-methods study of feedback modes in EFL writing. *Language Learning and Technology*, 24(1), 107–128.
- Berg, B. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8(3), 215–241. [https://doi.org/10.1016/S1060-3743\(99\)80115-5](https://doi.org/10.1016/S1060-3743(99)80115-5)
- Biber, D., Nekrasova, T., & Horn, B. (2011). The Effectiveness of Feedback for L1-English L2-Writing Development: A Meta-Analysis. *ETS Research Report Series*, 2011(1), i–99. <https://doi.org/10.1002/j.2333-8504.2011.tb02241.x>
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Wiley & Sons.
- Chen, L. S., & Renandya, W. A. (2020). Efficacy of written corrective feedback in writing instruction: A meta-analysis. *TESJ-EJ*, 24(3), 1–26.
- Chen, T. (2016). Technology-supported peer feedback in ESL/EFL writing classes: A research synthesis. *Computer Assisted Language Learning*, 29(2), 365–397. <https://doi.org/10.1080/09588221.2014.960942>
- Cheng, G. (2017). The impact of online automated feedback on students' reflective journal writing in an EFL course. *Internet and Higher Education*, 34, 18–27. <https://doi.org/10.1016/j.iheduc.2017.04.002>
- Cheng, K.-H., Liang, J.-C., & Tsai, C.-C. (2015). Examining the role of feedback messages in undergraduate students' writing performance during an online peer assessment activity. *Internet and Higher Education*, 25, 78–84. <https://doi.org/10.1016/j.iheduc.2015.02.001>
- Ciftci, H., & Kocoglu, Z. (2012). Effects of peer e-feedback on Turkish EFL students' writing performance. *Journal of Educational Computing Research*, 46(1), 61–84. <https://doi.org/10.2190/EC.46.1.c>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates.
- Cumming, A., Kantor, R., Powers, D., Santos, T., & Taylor, C. (2000). *TOEFL 2000 writing framework: A working paper*. <https://www.ets.org/Media/Research/pdf/RM-00-05.pdf>
- DiGiovanni, E., & Nagaswami, G. (2001). Online peer review: An alternative to face-to-face? *ELT Journal*, 55(3), 263–272. <https://doi.org/10.1093/elt/55.3.263>
- Elola, I., & Oskoz, A. (2010). Collaborative writing: Fostering foreign language and writing conventions development. *Language Learning & Technology*, 14(3), 51–71.
- Elola, I., & Oskoz, A. (2017). Writing with 21st century social tools in the L2 classroom: New literacies, genres, and writing practices. *Journal of Second Language Writing*, 36, 52–60. <https://doi.org/10.1016/j.jslw.2017.04.002>
- Ene, E., & Upton, T. A. (2018). Synchronous and asynchronous teacher electronic feedback and learner uptake in ESL composition. *Journal of Second Language Writing*, 41, 1–13.
- Ertmer, P. A., Richardson, J. C., Belland, B., Camin, D., Connolly, P., Coulthard, G., Lei, K., & Mong, C. (2007). Using peer feedback to enhance the quality of student online postings: An exploratory study. *Journal of Computer-Mediated Communication*, 12, 78–99. <https://doi.org/10.1111/j.1083-6101.2007.00331.x>
- Ge, Z.-G. (2011). Exploring e-learners' perceptions of net-based peer-reviewed English writing. *International Journal of Computer-Supported Collaborative Learning*, 6(1), 75–91. <https://doi.org/10.1007/s11412-010-9103-7>
- Guasch, T., Espasa, A., Alvarez, I. M., & Kirschner, P. A. (2013). Effects of feedback on collaborative writing in an online learning environment. *Distance Education*, 34(3), 324–338. <https://doi.org/10.1080/01587919.2013.835772>
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461–473. <https://doi.org/10.1093/applin/amp048>
- Huisman, B., Saab, N., Broek, P., & v. d., & Driel, J. v. (2019). The impact of formative peer feedback on higher education students' academic writing: A meta-analysis. *Assessment & Evaluation in Higher Education*, 44(6), 863–880. <https://doi.org/10.1080/02602938.2018.1545896>
- Hyland, K., & Hyland, F. (2019). *Feedback in second language writing: Contexts and issues*. Cambridge University Press.
- Johnson, W. F., Stellmack, M. A., & Barthel, A. L. (2019). Format of instructor feedback on student writing assignments affects feedback quality and student performance. *Teaching of Psychology*, 46(1), 16–21. <https://doi.org/10.1177/0098628318816131>
- Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal*, 99, 1–18. <https://doi.org/10.1111/modl.12189>
- Kellogg, R. T., Whiteford, A. P., & Quinlan, T. (2010). Does automated feedback help students learn to write? *Journal of Educational Computing Research*, 42(2), 173–196. <https://doi.org/10.2190/EC.42.2.c>
- Kourbani, V. (2017). Writing center asynchronous/synchronous online feedback: The relationship between e-feedback and its impact on student satisfaction, learning, and textual revision. In R. Aaron & R. K. S. Amant (Eds.), *Thinking globally, composing locally: Rethinking online writing in the age of the global Internet* (pp. 233–256). Utah State University Press.
- Latifi, S., Noroozi, O., Hatami, J., & Biemans, H. J. A. (2019). How does online peer feedback improve argumentative essay writing and learning? *Innovations in Education and Teaching International*. <https://doi.org/10.1080/14703297.2019.1687005>
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning & Technology*, 60, 309–365. <https://doi.org/10.1111/j.1467-9922.2010.00561.x>
- Link, S., Mehrzad, M., & Rahimi, M. (2020). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2020.1743323>
- Liou, H. C., & Peng, Z. Y. (2009). Training effects on computer-mediated peer review. *System*, 37(3), 514–525. <https://doi.org/10.1016/j.system.2009.01.005>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. London: SAGE Publications, Inc.
- Ma, Q. (2019). Examining the role of inter-group peer online feedback on wiki writing in an EAP context. *Computer Assisted Language Learning*, 33(3), 197–216. <https://doi.org/10.1080/09588221.2018.1556703>
- McCabe, J., Doerflinger, A., & Fox, R. (2011). Student and faculty perceptions of E-feedback. *Teaching of Psychology*, 38, 173–179. <https://doi.org/10.1177/0098628311411794>
- McGrath, A., & Atkinson-Leadbetter, K. (2016). Instructor comments on student writing: Learner response to electronic written

- feedback. *Transformative Dialogues: Teaching & Learning Journal*, 8, 1–16.
- Min, H. T. (2005). Training students to become successful peer reviewers. *System*, 33(2), 293–308. <https://doi.org/10.1016/j.system.2004.11.003>
- 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. <https://doi.org/10.1016/j.jslw.2006.01.003>
- Morch, A. I., Engeness, I., Cheng, V. C., Cheung, W. K., & Wong, K. C. (2017). EssayCritic: Writing to learn with a knowledge-based design critiquing system. *Educational Technology & Society*, 20(2), 213–223.
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. V. Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 125–143). Erlbaum.
- Noroozi, O., Biemans, H., & Mulder, M. (2016). Relations between scripted online peer feedback processes and quality of written argumentative essay. *Internet and Higher Education*, 31, 20–31. <https://doi.org/10.1016/j.iheduc.2016.05.002>
- Noroozi, O., & Hatami, J. (2019). The effects of online peer feedback and epistemic beliefs on students' argumentation-based learning. *Innovations in Education and Teaching International*, 56(5), 548–557. <https://doi.org/10.1080/14703297.2018.1431143>
- Norris, J. M., & Ortega, L. (2006). *Synthesizing Research on Language Learning and Teaching*. Benjamin.
- Novakovich, J. (2016). Fostering critical thinking and reflection through blog-mediated peer feedback. *Journal of Computer Assisted Learning*, 32(1), 16–30. <https://doi.org/10.1111/jcal.12114>
- Onrubia, J., & Engel, A. (2009). Strategies for collaborative writing and phases of knowledge construction in CSCL environments. *Computers & Education*, 53, 1256–1265. <https://doi.org/10.1016/j.compedu.2009.06.008>
- Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics*, 30, 85–110. <https://doi.org/10.1017/S0267190510000115>
- Pae, J.-K. (2011). Collaborative writing versus individual writing: Fluency, accuracy, complexity, and essay score. *Multimedia-Assisted Language Learning*, 14(1), 121–148. <https://doi.org/10.15702/mall.2011.14.1.121>
- Pham, T. N., Lin, M., Trinh, V. Q., & Bui, L. T. P. (2020). Electronic peer feedback, EFL academic writing and reflective thinking: Evidence from a Confucian context. *Sage Open*. <https://doi.org/10.1177/2158244020914554>
- Pham, V. P. H. (2019). The effects of lecturer's model e-comments on graduate students' peer e-comments and writing revision. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2019.1609521>
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Sage Publications.
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Wiley.
- Saricaoglu, A. (2019). The impact of automated feedback on L2 learners' written causal explanations. *ReCALL*, 31(2), 189–203. <https://doi.org/10.1017/S095834401800006X>
- Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, 9(3), 259–291. [https://doi.org/10.1016/S1060-3743\(00\)00028-X](https://doi.org/10.1016/S1060-3743(00)00028-X)
- Schultz, J. M. (2000). Computers and collaborative writing in the foreign language curriculum. In M. Warschauer & R. Kern (Eds.), *Network-based language teaching: Concepts and practice* (pp. 121–150). Cambridge University Press.
- Shang, H.-F. (2019). Exploring online peer feedback and automated corrective feedback on EFL writing performance. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2019.1629601>
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in second language. *Journal of Second Language Writing*, 20(4), 286–305. <https://doi.org/10.1016/j.jslw.2011.05.010>
- Stevenson, M., & Phakiti, A. (2019). Automated feedback and second language writing. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing* (pp. 125–142). Cambridge University Press.
- Storch, N. (2013). *Collaborative writing in L2 classrooms*. Multilingual Matters.
- Tai, H.-C., Lin, W.-C., & Yang, S. C. (2015). Exploring the effects of peer review and teachers' corrective feedback on EFL students' online writing performance. *Journal of Educational Computing Research*, 53(2), 284–309. <https://doi.org/10.1177/0735633115597490>
- Tolosa, C., East, M., & Villers, H. (2013). Online peer feedback in beginners' writing tasks. *IALLT Journal of Language Learning Technologies*, 43(1), 1–24. <https://doi.org/10.17161/iallt.v43i1.8516>
- Tuzi, F. (2004). The impact of e-feedback on the revisions of L2 writers in an academic writing course. *Computers and Composition*, 21(2), 217–235. <https://doi.org/10.1016/j.compcom.2004.02.003>
- Vaezi, S., & Abbaspour, E. (2015). Asynchronous online peer written corrective feedback: Effects and affects. In M. Rahimi (Ed.), *Handbook of research on individual differences in computer-assisted language learning* (pp. 271–297). IGI Global.
- 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. <https://doi.org/10.1016/j.jslw.2006.09.004>
- Yang, Y.-F. (2016). Transforming and constructing academic knowledge through online peer feedback in summary writing. *Computer Assisted Language Learning*, 29(4), 683–702. <https://doi.org/10.1080/09588221.2015.1016440>
- Yang, Y.-F., & Meng, W.-T. (2013). The effects of online feedback training on students' text revision. *Language Learning & Technology*, 17(2), 220–238.
- Yu, S. L. (2020). Giving genre-based peer feedback in academic writing: Sources of knowledge and skills, difficulties and challenges. *Assessment & Evaluation in Higher Education*. <https://doi.org/10.1080/02602938.2020.1742872>
- Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102. <https://doi.org/10.1016/j.asw.2018.02.004>
- Zhu, M., Liu, O. L., & Lee, H.-S. (2020). The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing. *Computers & Education*, 143, 103668. <https://doi.org/10.1016/j.compedu.2019.103668>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.