

RESEARCH ARTICLE

Open Access



Turnitin peer feedback: controversial vs. non-controversial essays

Mohammed Abdullah Alharbi^{1*}  and Ali H. Al-Hoorie² 

* Correspondence: maalharbi@mu.edu.sa

¹Department of English, Faculty of Education, Majmaah University, Majmaah, 11952, Saudi Arabia
Full list of author information is available at the end of the article

Abstract

Although an important goal of learner peer feedback is promoting critical thinking, little attention has been paid to the nature of the topic, particularly whether it is controversial. In this article, we report on a study where 52 English majors were asked to comment on essays written by their peers using the PeerMark module of Turnitin. Half of the essays were about controversial topics in the Saudi society (e.g., women driving and banning cigarettes), whereas the other half were less controversial (e.g., importance of sleep and respecting parents). Our results showed that the participants provided significantly more critical, global comments on controversial essays. At the same time, this increase in global comments did not come at the expense of local, language-related comments—in that the participants did not provide significantly fewer language comments on controversial essays. The participants also reported favorable attitudes toward this task due to the convenience and anonymity of online feedback, thus allowing them to express their opinions more freely on controversial topics. We therefore concluded that utilizing an online platform that permits double-blind peer review on controversial essays seems to have the potential to stimulate critical thinking among language learners.

Keywords: Turnitin, Peer review, Controversial topics, Critical feedback, Online feedback

Introduction

For decades, peer feedback has attracted a great deal of interest for researchers in both first language and second language (L2) learning. This interest reflects the potential of peer feedback for fostering critical thinking (e.g., Lam, 2013; Li & Li, 2017; Min, 2006; Tsui & Ng, 2000), for the development of writing competence (Chen, 2014; Ferris & Hedgcock, 2005; Flower & Hayes, 1981; Ho & Savignon, 2007; Hyland & Hyland, 2019; Zamel, 1982), and for uncovering the processes and dynamics involved in L2 peer interaction (e.g., Chang, 2016a; Li & Li, 2017; Li & Li, 2018; Wu, Petit, & Chen, 2015). Some research has shown that peer feedback helps to engage learners in writing and revision effectively (e.g., Yang, Badger, & Yu, 2006) and helps enhance audience awareness and appreciation of what purpose a piece of writing should serve (e.g., Lee, 2015).



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Nevertheless, scholars have also reported challenges (e.g., class discipline, time management, student resistance) in peer feedback conducted in face-to-face settings (Ferris & Hedgcock, 2005; Kim, 2015; Leki, 1990; Liu & Hansen, 2002). Therefore, researchers have begun to investigate the use of technology to facilitate peer feedback such as wikis (e.g., Bradley, 2014), synchronous chatting (e.g., Chang, 2012; Liang, 2010), Turnitin-based peer feedback (Li & Li, 2017; Li & Li, 2018), and blogs (e.g., Chen, 2012), significantly widening the range of tools available. In this study, we examined the role of how controversial the writing topic was on the type of feedback given by peers through Turnitin. Although Turnitin is most famously known as a plagiarism detection tool for student essays (Buckley & Cowap, 2013; Penketh & Beaumont, 2014), its new tool, PeerMark, is specifically designed to facilitate online peer feedback.

Literature review

Peer feedback in L2 writing

Learner peer feedback (or peer review) refers to the process of collaboration between learners on improving a piece of writing. The writer may receive comments, orally or in writing, during the initial stages of brainstorming and outlining up to the very final stages when the writing is complete. According to a systematic review by Chang (2016b), previous research on peer feedback has focused on one of three major lines of research. One line of research examined the *perceptions and attitudes* of learners toward peer feedback as a pedagogic strategy. Another line of research examined the *process* of peer feedback and its different implementation procedures. A third line of research looked at the *product* of peer feedback and what learning outcomes are achieved from it.

In relation to product-focused peer feedback, some investigators examined how effective it can be in improving the learner's writing. Some research suggests that learners can provide constructive comments to their peers, and the majority of this feedback is subsequently incorporated to improve the quality of writing (Li, 2018). Product-focused peer feedback has also been classified in different ways (e.g., Chang, 2012; Li & Li, 2017). One classification that is helpful in shedding light on the impact of peer feedback is to distinguish between local and global comments. Global comments involve feedback on development of ideas, purpose, organization, argumentation, and identification of audience; local comments focus more narrowly on language-related issues such as vocabulary, phrasing, syntax, and mechanics.

Some research examining this global–local distinction has explored its manifestation in online versus face-to-face settings. An early study by Hewett (2000) suggested that face-to-face feedback may place more emphasis on global areas whereas online feedback may be more concerned with local areas (see also Tuzi, 2004). However, other researchers argued that this pattern may not be inherently related to online versus face-to-face settings. That is, it is possible that some learners may overemphasize local rather than global areas simply due to lack of experience in providing feedback (Tsui & Ng, 2000) or to more familiarity with pen-and-paper-based feedback compared to software-based feedback (e.g., Bradley, 2014). Other variables affecting feedback type include cultural differences and deficient L2 rhetorical schemes (e.g., Ferris & Hedgcock, 2005). Liu and Sadler (2003) additionally argued that differences could emerge as a

result of mode of interaction (i.e., synchronous vs. asynchronous). Deliberate training and modeling can also increase the quantity and quality of global comments (e.g., Chang, 2015; Min, 2005, 2006; Rahimi, 2013).

Finally, some research suggests that online peer feedback might be superior to face-to-face feedback in certain respects. Online peer feedback tends to generate a larger number of comments (Chang, 2012), and these comments tend to be more revision-oriented, thus helping learners improve their essays. The convenience that technology provides also seems to contribute to the quality of comments compared to face-to-face peer feedback (Sengupta, 2001; Tuzi, 2004). Unlike face-to-face feedback, online peer feedback seems to allow the learner to read essays more carefully at their own pace and then formulate more quality feedback. Online peer feedback has additionally been found to promote more participation of learners (Chang, 2016b; Guardado & Shi, 2007) due to the comparatively less threatening situation for learners who have concerns about their English speaking proficiency and consequently prefer to remain silent in a conventional classroom setting (Liu & Sadler, 2003). Moreover, online peer feedback may reduce stress, especially when typed-up comments can be made anonymous (e.g., Bradley, 2014; Coté, 2014; Rourke, Mendelsohn, Coleman, & Allen, 2008).

From the above review, it is clear that while research to date has investigated various aspects related to peer feedback, little attention has been paid to the writing topic itself. Some topics have more personal meaning and value, potentially eliciting a different pattern of feedback as a result. Other topics might have controversial connotations (e.g., political, religious), and so learners might engage differently with peer feedback. Peer feedback research has generally tended to avoid such controversial topics perhaps due to the sensitivity of this approach. However, the anonymity of online feedback might encourage deeper engagement in these topics, thus plausibly stimulating critical thinking.

Critical thinking in L2 writing

In our context, critical thinking can be defined as “students’ appropriate reasoning, questioning, and critical evaluation, when receiving new information or applying previous knowledge to new situations” (Liu, Lin, Chiu, & Yuan, 2001, p. 248). Because of its potential value, the topic of critical thinking has received some attention in L2 writing research (e.g., Floyd, 2011; Indah, 2017; Miri & Azizi, 2018). Critical thinking seems to be crucial to success in university-level writing courses (Cheung, Rudowicz, Kwan, & Yue, 2002; Floyd, 2011; Moore, 2004; Phillips & Bond, 2004). As Weigle (2002) states, “Writing and critical thinking are seen as closely linked, and expertise in writing is seen as an indication that students have mastered the cognitive skills required for university work” (p. 5). Critical thinking is a challenge for L2 learners not only because they need to utilize higher-order thinking skills, but also because they have to do so in a second language (Floyd, 2011).

A number of experimental studies have shown that introducing critical thinking activities to writing courses improves learners’ writing skills. In one study by Khodabakhsh, Jahandar, and Khodabandehlou (2013), the researchers examined the impact of critical thinking tasks on paragraph writing ability of L2 learners. The experimental participants, who tackled critical thinking tasks, improved their writing skills in

comparison to the control participants. Likewise, Miri and Azizi (2018) investigated the impact of teaching critical thinking skills on L2 learners' writing ability. The experimental group received more instruction and practice on critical thinking techniques. The researchers found that these critical thinking techniques significantly improved the experimental learners' writing skills.

Engagement with peer review, particularly global feedback, seems to be an activity that has the potential to promote critical thinking. This process has been described "a critical thinking spillover effect" (DeLoach & Greenlaw, 2005, p. 150) that is augmented through the use of technology. Especially when the topic of discussion is controversial, the L2 learner has the opportunity to reflect on a peer's argumentation and critique it. To do that, the learner has to conceptualize, analyze, synthesize, and evaluate the argument presented by the essay writer. The learner is additionally expected to sustain a level of critical thinking, involving higher-order thinking and reasoning skills, even when a peer's writing is not perfectly intelligible and when asking for clarification is not an option. Besides, since the learner is aware that the essay writer is a (mortal) peer, rather than an expert or another authority figure, this may further encourage him/her to think more critically (see also Lynch, McNamara, & Seery, 2012; McMahon, 2010; Yang, 2016, for similar arguments).

Findings by Li and Li (2017) provide some evidence that online peer feedback can stimulate critical thinking when the writing topic revolves around a controversial issue. In that study, the researchers gave 13 L2 learners two tasks: summary–response and argumentative. The summary–response task elicited slightly more local comments than global comments. When it came to the argumentative task, however, the global comments elicited were more than double the number of local comments. This pattern could be due to the fact that the argumentative task was about "an up-to-date contentious topic" (Li & Li, 2017, p. 24), thus encouraging peers to bring in their own attitudes, beliefs, and convictions. Thanks to the convenience technology affords, "the possibility to access all papers and all reviews seemed to have played an important role in sharpening their critical thinking skills by addressing global issues" (Li & Li, 2017, p. 34). Therefore, it seems plausible that engaging learners in peer review of controversial essays can be a useful strategy to stimulate global feedback.

Turnitin

Turnitin has primarily been used for originality check and plagiarism detection in a written text (Rolfe, 2011). However, Turnitin also contains two other modules, namely GradeMark and PeerMark, to help students monitor their learning and gain support in academic writing (Li, 2018; Li & Li, 2017, 2018). Some researchers have examined instructor feedback on student writing via the GradeMark tool of Turnitin, where teachers provide electronic comments on each student's paper submitted to Turnitin. For instance, Buckley and Cowap (2013) explored 11 university instructors' perspectives on the use of the GradeMark feature. The participants mentioned a number of strengths, most notably the ease and speed of marking certain assignment formats. Buckley and Cowap's findings are in line with those by Henderson (2008), who noted that providing online feedback through Turnitin automates parts of the process. This for example allows the marker to avoid having to handwrite the same comments

repeatedly to different learners. Instructor comments on GradeMark deal with various issues including critical aspects of assignments, referencing, grammar, and writing mechanics (Reed, Watmough, & Duvall, 2015).

PeerMark is another feature of Turnitin that has only recently attracted writing researchers' attention (see Fig. 1). The primary function of PeerMark is for learners to read and comment on each other's essays. When students submit their essays through Turnitin, the instructor can either start grading and giving feedback on each essay, or forward these essays to other students in the class to read and evaluate. Both essay writers and peer reviewers can either be attributed or anonymized by the class instructor. The instructor can also assign a certain number of essays for each student to review by a deadline, after which reviews can no longer be written, completed, or edited. Upon completing a review, feedback becomes immediately available to the essay writer.

Peer reviewers can read, review, score, and evaluate each essay. To do that, peer reviewers have three main tools at their disposal, namely commenting tools, composition marks, and direct questions. Commenting tools allow peer reviewers to highlight part of a text and insert a comment on it (see Fig. 2). The second tool, composition marks, offers peer reviewers a set of symbols representing common issues found in student writing. Examples of these symbols include word choice, spelling errors, and run-on sentences (see Fig. 3). Because these issues are common, composition marks allow peer reviewers to save time by dragging the relevant symbol and dropping it where the issue appears in the text. Finally, direct questions are a set of specific questions about the quality of the essay set in advance by the class instructor. Peer reviewers are expected to answer these questions, which can be free response or scale questions, as they evaluate the essay.

Our knowledge about PeerMark and how students utilize it is still limited since only a handful of studies have examined it to date. In a study by Li and Li (2017), the authors found that the majority of students' peer feedback was revision-oriented on both local areas (e.g., grammar, mechanics, vocabulary, and format) and global areas (e.g., organization, content, and referencing). The participants also evaluated their experience with this tool favorably as reflected both by questionnaires and interviews. In another study by Li and Li (2018), both the class instructor and the students found this tool

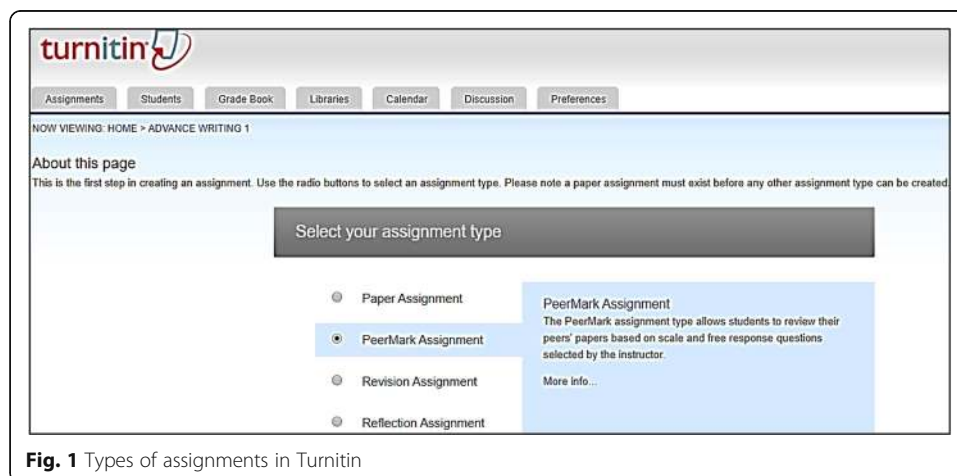


Fig. 1 Types of assignments in Turnitin



Fig. 2 Sample of a peer reviewer's comments in PeerMark

helpful in shifting the attention of students from local to global issues and from holistic advice to specific suggestions, as well as facilitating classroom management (see also Li, 2018).

The present study

As we explained above, although some research has been conducted on peer feedback and critical thinking, the type of essay—particularly whether it is dealing with a controversial topic—has not been systematically investigated. The primary aim of this study was therefore to examine the potential of controversial essays in stimulating global comments. Giving global feedback transcends “superficial” language-related issues, requiring a deeper-level understanding of the writer’s argument, the evidence provided to support it, and how this evidence is organized (Li & Li, 2017; Lynch et al., 2012; McMahon, 2010; Yang, 2016).

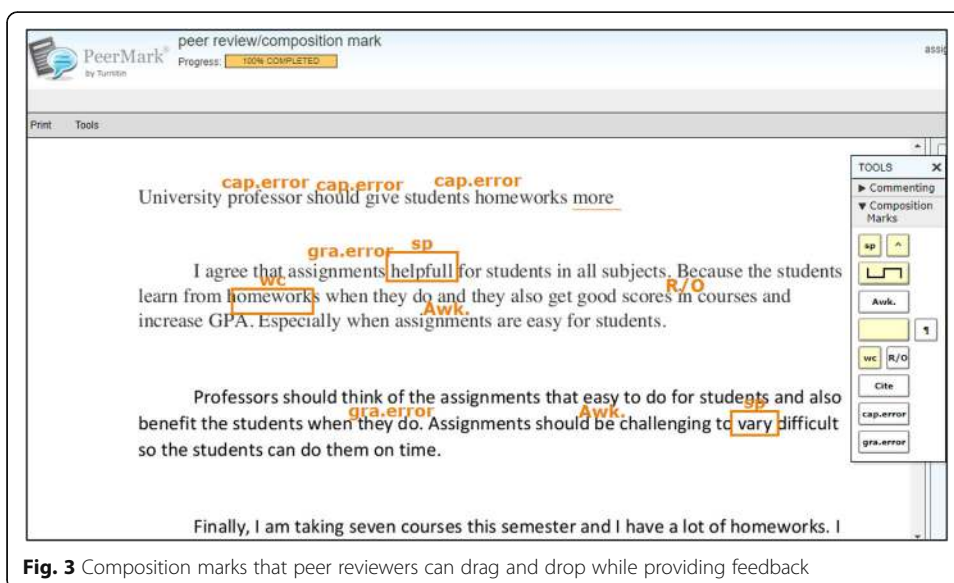


Fig. 3 Composition marks that peer reviewers can drag and drop while providing feedback

At the same time, because the task is in the L2, it is not clear whether focus on global issues in controversial essays distracts from local issues. That is, it is possible that, in controversial essays, local areas suffer in that learners receive fewer comments on local issues from their peers.

Finally, we wanted to investigate the subjective experience of learners during this activity. That is, even if learners objectively give more global comments on controversial essays, the potentially sensitive nature of these topics might hinder full engagement with them. Nevertheless, considering that all tasks can be performed in a double-blind fashion, we expected that the participants would not feel uneasy about engaging with these tasks.

To summarize, we posed the following three hypotheses:

Hypothesis 1: Students give more global comments in controversial essays, compared with non-controversial essays.

Hypothesis 2: Students give fewer local comments in controversial essays, compared with non-controversial essays.

Hypothesis 3: Students do not feel uneasy about engaging with essays on controversial topics, considering the double-blind nature of the process.

Addressing these issues has important pedagogical implications. If local comments do become scarce in controversial essays, this may suggest that this approach is not suitable for lessons with a central focus on improving local areas. Conversely, if the number of global comments in controversial essays turns out to be comparable to that in non-controversial essays, then instructors won't need to look for controversial topics to elicit global feedback.

Method

Participants

One class of 52 university students participated in this study. The participants (aged 20–22) were male English majors at a major all-male Saudi university. They had been admitted to the English department based on their IELTS exam scores, which is one of the program's entrance requirements. The department requires a score of at least 4.0 on the IELTS to be admitted to the program.

The class in which this study was conducted was an advanced writing course offered by the English department. The program the participants were enrolled in consists of eight levels of English (each level is one semester long). The participants were at Level 4, which is the last level of the second year. The students were therefore considered to be at around an intermediate level in writing by this point.

Instrument

The participants used Turnitin to write essays and then obtain feedback from their peers (see Procedure). The feedback was given using the PeerMark functionality in Turnitin, which allows peers to make comments and forward them to the original authors. Toward the end of the study, the participants completed an anonymous online questionnaire adapted from Li and Li (2017) evaluating their experience using Turnitin.

This questionnaire consisted of 12 items rated on a 5-point Likert scale where a higher score indicates a more favorable attitude. The participants additionally answered open-ended questions about their overall impression about the use of Turnitin for this task. Because the participants were English majors, the questionnaire was administered in English. The questionnaire was completed in class, and the instructor was present to answer any queries about it.

Procedure

The first step was to decide on controversial and non-controversial topics. We created a list of 20 statements dealing with various social topics related to the Saudi society and asked a group of students not participating in this study to rate the extent to which they agreed or disagreed with each statement. These students were 44 in total and came from the same institution as those participating in this study. We then selected six statements that received a high level of agreement (around 90%) to serve as prompts for non-controversial essays. Six other statements that received mixed responses—with a roughly equivalent amount of agreement, disagreement, and undecidedness—were selected for controversial essays (see Table 1 for the full list of prompts used in this study). This was an important step to ensure that controversial and non-controversial prompts were perceived as such by a sample from the target population.

Second, we randomly selected 24 participants from the class participating in this study to write a short essay using this list of prompts. This process generated two essays about each prompt in Table 1. This pool of 24 essays constituted an authentic basis for the present study.

Third, the class instructor, who was the first author of the present paper, introduced the PeerMark tool to the participants and step-by-step trained them on how to access, read, and comment on these essays. The instructor also made it clear to the participants that comments can either be on local or on global areas, illustrating each with a variety of examples.

Fourth, we forwarded the essays to the participants using the PeerMark functionality. To ensure that controversial and non-controversial essays received equivalent exposure, we randomly assigned the essays to the participants after anonymizing both essay writers and peer reviewers. The participants read and commented on the essays within 2 weeks outside of class time. On average, each participant commented on between 8

Table 1 The list of controversial and non-controversial prompts used in this study

Controversial Prompts	Non-controversial Prompts
Schoolchildren should have the choice to study English or not.	Learning English is important nowadays.
Cigarettes should be illegal in Saudi Arabia.	Smoking is bad for health.
English should be taught by native speakers only.	The Internet has become important in daily life.
Watching sports is a waste of time.	Getting enough sleep is important for health.
University professors should give students more homework.	Reading helps develop the mind.
Saudi women are better drivers than Saudi men. ^a	It is important to respect your parents.

^a This topic attracted considerable attention at the time of this study due to a royal decree lifting a decades-old ban on women driving

and 9 essays. Each essay received a generally comparable number of comments from each learner who reviewed it (typically 0–4), while the highest number of comments by one learner on a single essay was 10 (all language-related comments on the non-controversial prompt *Learning English is important nowadays*).

We obtained IRB approval from the Deanship of Scientific Research at Majmaah University. All participants signed an informed consent form indicating their desire to take part in the study. The participants received no compensation for participation. Throughout, the participants were treated according to APA ethical guidelines.

Data analysis

We coded the comments to determine whether each represented a local or a global issue. A local area involved commenting on issues related to vocabulary, grammar, mechanics, and format. Global comments related to the essay content, its organization, and the logic and evidence used to support the writer's position. At first, two experts with PhDs in Applied Linguistics coded 10% of the comments, Cohen's $\kappa = .97$, $p < .001$. Subsequently, one coder coded the remaining comments. In very few cases, one comment contained both local- and global-related aspects (e.g., disagreeing with the argument and then pointing out a problem with word usage). In these cases, the comment was divided into two comments: local and global.

In terms of the evaluation survey, the questionnaire scale was analyzed using a one-sample *t*-test (see Al-Hoorie & Vitta, 2019) to examine whether the overall mean was significantly different from middle point of 3.0, which represents indifference. To ensure trustworthiness, two trained external raters conducted a line-by-line open coding of the open-ended survey responses (Glaser & Strauss, 1967; Saldaña, 2016). They read and reread responses to ensure validity and avoid bias in the analysis process. They eventually identified three themes emerging from the data, two related to favorable aspects of the Turnitin experience and the other related to areas needing improvement in this platform to make the activity more effective.

Results

Table 2 presents descriptive statistics of the results. When the study concluded, Peer-Mark generated a total of 618 comments. On average, a participant made 11.88 comments. The majority of these comments (68.1%) were local comments, while fewer than a third were global comments. This indicates that for every two local comments, there was only about one global comment.

The following are examples of local comments made by the participants:

- You said first "The internet has become one of the most important thing in our time" well you made a mistake here by saying "Thing" it should be (Things) because you are talking about the internet in general so yeah. After that the sentence continue by saying "And the internet has pros and cons..." since it's a continuation for the previous sentence you can not write the first word as a capital word so it should have been "and"
- you should write cities with capital letter at first

Table 2 Descriptive statistics of comments generated by the end of the study

	<i>n</i>	<i>M</i>	<i>SD</i>
Overall comments	618	11.88	7.04
Local comments	421 (68.1%)	8.10	7.77
Global comments	197 (31.9%)	3.39	3.39

- you should put the title in the middle of the paper also you should make space in the beginning of the first line of the paragraph. Also when you finished from the sentence put full stop.

The following are examples of global comments:

- I agree with you on how homework teaches us, but more homework?! that sounds horrible. Some of the professors dont bother about how the students have a lot of things to do, like they think that the students have nothing to do or to study. to put the student under the pressure, no matter what he\she is human have limits, and we are human. More homewrk will stress the student and that is not healthy physical or mental
- I understand your whole idea in general though i got to say first, you said that cigarette is one of the main cause of air pollution. I disagree with you here because its not one of the main reason of it, factories and cars are the main reason
- I agree on how bad to watch tv all the time and damaging your eyes and stuff, but i do not agree with you about how watching sports is a waste of time. Its waste of time but the way you said it it sounded like you are holding something against sports. Sports are fun to watch like anything else, watching movies or tv shows or playing games. They don't have to be educational sometimes, you are just having fun and taking it easy. But even sports can teach you things like skills or something, or even it become the reason you want to be a player. Unhealthy addiction is not only on watching spotsrs, there is a lot of fun stuff if you were into them too much it will be unhealthy. So every fun thing has its limits.

Types of comments on controversial vs. non-controversial essays

To test Hypotheses 1 and 2, we compared the comments coded as either local or global in relation to whether they belonged to a controversial or non-controversial essay. A paired sample *t*-test showed the participants gave, on average, significantly more global comments on controversial essays compared with non-controversial essays. This finding provides support for Hypothesis 1, indicating that learners give more global comments on controversial essays.

In contrast, the number of local comments was not significantly different between controversial and non-controversial essays. This does not support Hypothesis 2. That is, this pattern provides no evidence that local feedback decreases when the essay topic is controversial. Table 3 presents these results.

Subjective experience during the task

We analyzed the survey responses quantitatively and qualitatively. Starting with the quantitative analysis, preliminary analysis showed that the scale was unidimensional

(through inspection of the scree plot) and had an adequate level of reliability, $\alpha = .87$. Descriptive statistics (see Table 4) showed that all items had a mean larger than 3.0, the middle point, indicating that the participants on average tended to agree with these statements ($M = 3.72$, $SD = 0.61$). Supporting this impression, a one-sample t -test showed that the mean of this scale was significantly higher than 3.0, $t(51) = 8.61$, $p < .001$, $d = 1.18$. These results suggest that the participants exhibited favorable attitudes toward this task.

As for the qualitative part, the participants provided open-ended responses about their experience with this activity. We analyzed these responses in order to obtain a deeper understanding of the participants' subjective experience while giving and receiving peer feedback on controversial essays. Analysis of the data uncovered three major themes, two concerning positive aspects of Turnitin and a third theme pointing to possible room for improvement.

The first major theme was the convenience Turnitin afforded the participants. Most of the participants had no prior experience with such an online peer feedback tool. They found the tool relatively easy to access and to use whether on or off campus. For example, one participant described Turnitin as *"easy for using, helpful for the students, easy to access, can use in your home or college or anywhere."* This convenience encouraged some of them to read and comment on more essays than they would have offline. Another participant commented that *"I like it because I use my phone after class in break and sometimes in park when I am free I read my friends feedback and correct my essay."* This gave them additional opportunities to learn not only from receiving feedback but also from giving it to their peers. Since all feedback is stored on the platform, the participants were also able to retrieve it at a later time for the purpose of further study.

A second major theme emerging from the analysis had to do with the anonymity of comments. The participants appreciated the option of making the whole peer review process double-blind. The anonymity encouraged the participants to be more honest and give critical comments especially on controversial and culturally sensitive topics. As one participant put it, it is *"like talking to a computer. You don't have to care about its reaction."* Similarly, comment receivers did not feel embarrassed or lose face when their peers pointed out flaws in their arguments. This in turn could promote critical thinking for both feedback givers and receivers.

Interestingly, despite the sensitive nature of some of the essays, no participant raised concerns about this aspect. Instead, they reported being comfortable reading and commenting on an essay containing arguments with which they disagreed when the process was double-blind. They found in this process an opportunity, on the one hand, to think critically and give thoughtful comments at their own pace and, on the other hand, to

Table 3 Paired sample t -test results for the two types of comments on controversial and non-controversial essays

	Essay	$n(\%)$	M	SD	$r(p)$	$t(df)$	p	d
Global Comments	Controversial	119(19.3)	2.29	2.35	.27(.05)	2.45(51)	.018	0.415
	Non-Controversial	75(12.1)	1.44	1.69				
Local Comments	Controversial	210(32.5)	3.87	3.77	.62(< .001)	0.78(51)	.439	0.096
	Non-Controversial	223(36.1)	4.29	4.88				

Table 4 Means and standard deviation of the questionnaire items (adapted from Li & Li, 2017)

Item		<i>M</i>	<i>SD</i>
1.	I enjoyed using Turnitin for peer feedback.	3.63	1.05
2.	My experience with Turnitin-based peer feedback was positive.	3.69	0.98
3.	I found Turnitin's originality report on my assignments helpful.	3.79	0.85
4.	I used Turnitin's originality report to revise my work.	3.62	1.07
5.	Turnitin has helped me become better at citing sources.	3.81	0.86
6.	I found the peer feedback I received on Turnitin helpful.	3.75	0.90
7.	I found the PeerMark questions helpful for my peer feedback.	3.75	0.99
8.	I enjoyed using PeerMark commenting tools to provide peer feedback.	3.60	1.05
9.	I enjoyed using composition marks to provide peer feedback.	3.69	0.97
10.	I prefer conducting peer feedback using Turnitin to doing face-to-face peer feedback.	3.69	1.15
11.	I incorporated Turnitin peer feedback into my revision.	3.77	0.67
12.	Peer feedback using Turnitin improves my writing skills.	3.90	0.80

reflect objectively about a position with which they disagree without having to feel defensive.

With regard to areas needing improvement, the most common theme emerging from the analysis was technical glitches while using Turnitin. Some participants reported occasional difficulties with submitting their work, and sometimes with logging in. As one participant explained, “*when I clicked on submit, the website closed and I returned to the main webpage.*” This was a frustrating experience for the participants who had encountered this issue, as they had to log in and type up their feedback all over again. Most participants used the website platform for this study, while only a few used the Turnitin smartphone app.

Other participants suggested introducing additional features to Turnitin. One suggestion was giving the reviewer the chance to submit a “*voice comment,*” as this option would allow the learner to additionally practice their speaking skill. Another suggestion was allowing peer reviewers access to perform plagiarism check themselves. In this way, the peers can put on their instructor hat and investigate whether the text was plagiarized and whether the original text was adequately paraphrased and cited by the essay writer. This activity may help L2 learners improve their writing skills by examining how their peers handle, paraphrase, quote, and cite sources—a process that tends to be performed by learners in an individual manner that is typically invisible to peers. Finally, one participant suggested implementing plagiarism check on the comments themselves as well. This might be relevant to more extended comments, especially in cases where learners are evaluated on the quality of their review. It would appear that this adding such features should not be technically challenging consider that Turnitin is originally a plagiarism detection tool.

Discussion

The primary purpose of this study was to examine the effectiveness of controversial topics in stimulating global comments from peer reviewers. More specifically, we hypothesized that asking learners to comment on controversial essays would result in more global comments compared with non-controversial essays (Hypothesis 1). At the same time, we also sought to find out whether attention to global issues in controversial

essays would result in a decrease in local comments (Hypothesis 2). The third purpose of this study was to explore the participants' subjective evaluation of this task, which we expected to be favorable considering its double-blind nature (Hypothesis 3).

In terms of Hypothesis 1, our results indicated that, indeed, asking learners to comment on controversial essays led to a significant increase in the number of global comments generated. This suggests that asking learners to comment on a controversial essay might be a useful strategy to stimulate critical, global comments. In terms of Hypothesis 2, our results do not provide evidence that this increase in global comments comes at the expense of local comments. The participants gave a comparable number of local comments, whether the essay was controversial or not.

Generally speaking, our results are consistent with previous research showing that learners tend to focus on local areas (Hewett, 2000; Tuzi, 2004), especially when these learners are left untrained (Chang, 2015; Min, 2005, 2006; Rahimi, 2013). The participants of the present study provided, overall, more comments on local than on global issues—only one third of all comments were global (see Table 2). Nevertheless, once we took into account the nature of the topic, interesting results emerged. Our findings showed that the pattern of feedback was directly related to the type of essay in question and whether it was dealing with a controversial topic. These findings seem in line with those by Li and Li (2017), whose participants provided more global comments on argumentative essays compared with summary–response essays. As Li and Li (2017) explained, the nature of summary–response tasks requires learners to devote a significant part of their feedback on summarizing the writer's point of view, thus leaving limited space for critical commentary.

To put it differently, not all topics are created equal. While some research on learner feedback has investigated whether learners focus on local or global areas, we consider it almost meaningless to ask this question out of the context of the topic in question. If learners consider a topic to be of neutral valence, they may consequently tend to focus more on local areas. In contrast, if the topic is “hot,” on a contentious subject, or holds personal and profound value, it is more likely that learners will engage more deeply with it. This level of affective engagement might also enhance cognitive engagement, potentially leading to better learning. Suggestive of this, our results showed that there was no significant reduction in local comments on controversial essays.

In relation to the participants' experience with the double-blind nature of the activity (Hypothesis 3), the participants exhibited favorable attitudes. Many participants stated that they preferred this “faceless” system. According to these participants, the ability to give anonymous feedback makes the process less stressful and, therefore, there is little incentive to engage in face-saving strategies at the expense of giving honest, constructive feedback to their peers. These findings support previous research showing the value of anonymizing online learner feedback (e.g., Bradley, 2014; Coté, 2014; Rourke et al., 2008). Other participants pointed out the convenience of giving feedback from home during their free time, instead of a situation where their peer is in the same room waiting (im) patiently for feedback. These results are in line with those by Li and Li (2017), who reported that their participants “unanimously” (p. 34) agreed that Turnitin was an effective and enjoyable tool to conduct learner peer review. In the present study, responses from open-ended questions indicated that the majority of our participants felt that online peer feedback gave them an opportunity to conveniently review the writing

of their peers and learn from each other. This also echoed Chen's (2012) results from blog-based peer feedback where technology not only facilitated peer feedback but also had a positive impact on the students' learning experience. All in all, while some research (e.g., Liu & Sadler, 2003) suggests that face-to-face feedback may be more effective due to the presence of nonverbal communication, our results suggest that the nature of the topic and certain technological features (such as anonymity and convenience) also play an important role in the amount and type of peer feedback given.

Pedagogical implications

The results of the present study have some pedagogical implications. Teachers wishing to stimulate critical thinking can utilize peer review on controversial topics. An obvious concern in applying this strategy is the risk involved in controversial and culturally sensitive topics. When a topic is too controversial, learners might be distracted and feedback might turn to pure airing of opinions that do not represent an objective evaluation of a peer's argument. It is possible that this type of engagement might not be as conducive to language learning or to the development of critical thinking skills. The prudent teacher should therefore aim to come up with topics that have an appropriate level of controversiality. When the instructor is a foreigner who is not very familiar with the host culture, it would be advisable to consult with other teachers, and with the students themselves, to find out what topics the learners are willing to discuss. These topics will most likely vary from culture to culture, and the topics we used in the present study might or might not be suitable in other contexts.

Using an online platform has additional pedagogical advantages. In the past, a platform like Turnitin used to be primarily a plagiarism detection tool. It was therefore a rather peripheral tool in the educational process, resorted to only prior to formal submission of essays. With a feature like PeerMark, however, technology serves a much more active role in the day-to-day learning process. This automation can free up some of the teacher's time, allowing them to focus on learners in need of individual assistance. The teacher will additionally have a record of what comments are given and who is (not) giving them. The teacher can subsequently then use these comments for classroom discussions.

Limitations and future directions

This study is not without limitations. While the primary goal of using controversial essays was to ultimately promote critical thinking, we measured global comments as a proxy for critical thinking. While a number of researchers have emphasized a strong link between critical thinking and global comments on essays written by peers (Li & Li, 2017; Lynch et al., 2012; McMahan, 2010; Yang, 2016), future research should investigate this link more directly by utilizing measures of critical thinking. However, it is likely that this type of research would need to be conducted over a longer period of time, as a sizeable increase in critical thinking might require more frequent engagement with peer review.

Another potential future direction is a more micro-analysis of peer review. We argued that not all essay prompts are the same in their effect on critical thinking. It is equally likely that not all strategies of engagement with peer review are effective in

stimulating critical thinking. Class teachers might need to deliberately draw the learners' attention to critical thinking strategies (e.g., Brookfield, 2012) for more effective implementation.

A further future direction is the effectiveness of group work. In our study, the participants worked individually, reading and commenting on essays by their peers. It would be interesting to examine how learners work as a group to evaluate an essay and give anonymous feedback on it.

Finally, some platforms allow the user to give oral feedback. Giving oral feedback might be more convenient to comment givers, but it might not be as convenient to comment receivers. Comments receivers, assuming that they understand the oral message, may find it time-consuming especially if they need to locate again a comment they came across previously. This limitation might be circumvented if the feedback giver's speech can be automatically transcribed.

Acknowledgements

Not applicable.

Authors' contributions

Both authors designed the study, conduct experiments, collect and analyzed the data and co-wrote the paper. The author(s) read and approved the final manuscript.

Authors' information

Dr. Mohammed Abdullah Alharbi is an assistant professor in the Department of English at Majmaah University, Saudi Arabia. His research interests include L2 writing instruction, computer-mediated communication, peer review, and peer interaction.

Ali H. Al-Hoorie is an assistant professor at the English Language and Preparatory Year Institute, Royal Commission for Jubail and Yanbu, Saudi Arabia. He completed his PhD in Applied Linguistics at the University of Nottingham under the supervision of Professors Zoltán Dörnyei and Norbert Schmitt. He also holds an MA in Social Science Data Analysis from Essex University. His research interests include motivation theory, research methodology, and complexity. His publications have appeared in various journals including *Language Learning*, *Modern Language Journal*, *Studies in Second Language Acquisition*, *ELT J*, *Language Teaching Research*, and *Learning and Individual Differences*. His books include *Research Methods for Complexity in Applied Linguistics* (Multilingual Matters, 2020, with Phil Hiver) and *Contemporary Language Motivation Theory: 60 Years Since Gardner and Lambert (1959)* (Multilingual Matters, 2020, coedited with Peter MacIntyre).

Funding

The authors would like to thank Deanship of Scientific Research at Majmaah University for supporting this work under project number: 1439-90.

Availability of data and materials

The data that support the findings of this study are available upon request.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of English, Faculty of Education, Majmaah University, Majmaah, 11952, Saudi Arabia. ²English Language and Preparatory Year Institute, Royal Commission for Jubail and Yanbu, Jubail, Saudi Arabia.

Received: 28 January 2020 Accepted: 18 March 2020

Published online: 13 May 2020

References

- Al-Hoorie, A. H., & Vitta, J. P. (2019). The seven sins of L2 research: A review of 30 journals' statistical quality and their CiteScore, SJR, SNIP, JCR impact factors. *Language Teaching Research*, 23(6), 727–744. <https://doi.org/10.1177/1362168818767191>.
- Bradley, L. (2014). Peer-reviewing in an intercultural wiki environment - student interaction and reflections. *Computers and Composition*, 34, 80–95. <https://doi.org/10.1016/j.compcom.2014.09.008>.
- Brookfield, S. D. (2012). *Teaching for critical thinking: Tools and techniques to help students question their assumptions*. San Francisco: Wiley.
- Buckley, E., & Cowap, L. (2013). An evaluation of the use of Turnitin for electronic submission and marking and as a formative feedback tool from an educator's perspective. *British Journal of Educational Technology*, 44(4), 562–570. <https://doi.org/10.1111/bjet.12054>.

- Chang, C. F. (2012). Peer review via three modes in an EFL writing course. *Computers and Composition*, 29(1), 63–78. <https://doi.org/10.1016/j.compcom.2012.01.001>.
- Chang, C. Y. H. (2015). Teacher modeling on EFL reviewers' audience-aware feedback and affectivity in L2 peer review. *Assessing Writing*, 25, 2–21. <https://doi.org/10.1016/j.asw.2015.04.001>.
- Chang, C. Y. H. (2016a). EFL reviewers' emoticon use in asynchronous computer-mediated peer response. *Computers and Composition*, 40, 1–18. <https://doi.org/10.1016/j.compcom.2016.03.008>.
- Chang, C. Y. H. (2016b). Two decades of research in L2 peer review. *Journal of Writing Research*, 8(1), 81–117. <https://doi.org/10.17239/jowr-2016.08.01.03>.
- Chen, K. T. C. (2012). Blog-based peer reviewing in EFL writing classrooms for Chinese speakers. *Computers and Composition*, 29(4), 280–291. <https://doi.org/10.1016/j.compcom.2012.09.004>.
- Chen, T. (2014). 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>.
- Cheung, C., Rudowicz, E., Kwan, A. S. F., & Yue, X. D. (2002). Assessing university students' general and specific critical thinking. *College Student Journal*, 36(4), 504–525.
- Coté, 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.
- Deloach, S. B., & Greenlaw, S. A. (2005). Do electronic discussions create critical thinking spillovers? *Contemporary Economic Policy*, 23(1), 149–163. <https://doi.org/10.1093/cep/byi012>.
- Ferris, D. R., & Hedgcock, J. S. (2005). *Teaching ESL composition: Purpose, process, and practice*. New York: Routledge.
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387.
- Floyd, C. B. (2011). Critical thinking in a second language. *Higher Education Research & Development*, 30(3), 289–302. <https://doi.org/10.1080/07294360.2010.501076>.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Piscataway: AldineTransaction.
- Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition*, 24(4), 443–461. <https://doi.org/10.1016/j.compcom.2007.03.002>.
- Henderson, P. (2008). Electronic grading and marking: A note on Turnitin's Grademark function. *History Australia*, 5(1), 11.1–11.2. <https://doi.org/10.2104/ha080011>.
- Hewett, B. L. (2000). Characteristics of interactive oral and computer-mediated peer group talk and its influence on revision. *Computers and Composition*, 17(3), 265–288. [https://doi.org/10.1016/S8755-4615\(00\)00035-9](https://doi.org/10.1016/S8755-4615(00)00035-9).
- Ho, M. C., & Savignon, S. J. (2007). Face-to-face and computer-mediated peer review in EFL writing. *CALICO Journal*, 24(2), 269–290.
- Hyland, K., & Hyland, F. (2019). Contexts and issues in feedback on L2 writing. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues*, (2nd ed., pp. 1–22). Cambridge: Cambridge University Press.
- Indah, R. N. (2017). Critical thinking, writing performance and topic familiarity of Indonesian EFL learners. *Journal of Language Teaching and Research*, 8(2), 229–236.
- Khodabakhsh, S., Jahandar, S., & Khodabandehlou, M. (2013). The impact of critical thinking tasks on paragraph writing ability of Iranian EFL learners. *Indian Journal of Fundamental and Applied Life Sciences*, 3(3), 639–648.
- Kim, S. H. (2015). Preparing English learners for effective peer review in the writers' workshop. *Reading Teacher*, 68(8), 599–603. <https://doi.org/10.1002/trtr.1358>.
- Lam, R. (2013). The relationship between assessment types and text revision. *ELT Journal*, 67(4), 446–458. <https://doi.org/10.1093/elt/cct034>.
- Lee, M. K. (2015). Peer feedback in second language writing: Investigating junior secondary students' perspectives on inter-feedback and intra-feedback. *System*, 55, 1–10. <https://doi.org/10.1016/j.system.2015.08.003>.
- Leki, I. (1990). Potential problems with peer responding in ESL writing classes. *CATESOL Journal*, 3(1), 5–19.
- Li, J., & Li, M. (2018). Turnitin and peer review in ESL academic writing classrooms. *Language Learning & Technology*, 22(1), 27–41. Doi:10.125/44576.
- Li, M. (2018). Online peer review using Turnitin PeerMark. *Journal of Response to Writing*, 4(2), 99–117.
- Li, M., & Li, J. (2017). Online peer review using Turnitin in first-year writing classes. *Computers and Composition*, 46, 21–38. <https://doi.org/10.1016/j.compcom.2017.09.001>.
- Liang, M. Y. (2010). Using synchronous online peer response groups in EFL writing: Revision-related discourse. *Language Learning and Technology*, 14(1), 45–64.
- Liu, E. Z. F., Lin, S. S., Chiu, C. H., & Yuan, S. M. (2001). Web-based peer review: The learner as both adapter and reviewer. *IEEE Transactions on Education*, 44(3), 246–251. <https://doi.org/10.1109/13.940995>.
- Liu, J., & Hansen, J. G. (2002). *Peer response in second language writing classrooms*. Ann Arbor: University of Michigan Press.
- Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, 2(3), 193–227. [https://doi.org/10.1016/S1475-1585\(03\)00025-0](https://doi.org/10.1016/S1475-1585(03)00025-0).
- Lynch, R., McNamara, P. M., & Seery, N. (2012). Promoting deep learning in a teacher education programme through self and peer-assessment and feedback. *European Journal of Teacher Education*, 35(2), 179–197. <https://doi.org/10.1080/02619768.2011.643396>.
- McMahon, T. (2010). Peer feedback in an undergraduate programme: Using action research to overcome students' reluctance to criticise. *Educational Action Research*, 18(2), 273–287. <https://doi.org/10.1080/09650791003741814>.
- Min, H.-T. (2005). Training students to become successful peer reviewers. *System*, 33, 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.jslwr.2006.01.003>.
- Miri, F., & Azizi, D. B. (2018). The effect of teaching critical thinking on Iranian EFL learners' essay writing. *Theory and Practice in Language Studies*, 8(5), 509–515.
- Moore, T. (2004). The critical thinking debate: How general are general thinking skills? *Higher Education Research & Development*, 23(1), 3–18. <https://doi.org/10.1080/0729436032000168469>.
- Penketh, C., & Beaumont, C. (2014). Turnitin said it wasn't happy: Can the regulatory discourse of plagiarism detection operate as a change artefact for writing development? *Innovations in Education and Teaching International*, 51(1), 95–104. <https://doi.org/10.1080/14703297.2013.796721>.

- Phillips, V., & Bond, C. (2004). Undergraduates' experiences of critical thinking. *Higher Education Research & Development*, 23(3), 277–294. <https://doi.org/10.1080/0729436042000235409>.
- Rahimi, M. (2013). Is training student reviewers worth its while? A study of how training influences the quality of students' feedback and writing. *Language Teaching Research*, 17, 67–89. <https://doi.org/10.1177/1362168812459151>.
- Reed, P., Watmough, S., & Duvall, P. (2015). Assessment analytics using Turnitin & Grademark in an undergraduate medical curriculum. *Journal of Perspectives in Applied Academic Practice*, 3(2), 92–108.
- Rolfe, V. (2011). Can Turnitin be used to provide instant formative feedback? *British Journal of Educational Technology*, 42(4), 701–710. <https://doi.org/10.1111/j.1467-8535.2010.01091.x>.
- Rourke, A. J., Mendelsohn, J., Coleman, K., & Allen, B. (2008). Did I mention it's anonymous? The triumphs and pitfalls of online peer review. In *Hello! Where are you in the landscape of educational technology? Proceedings ascilite Melbourne 2008* <http://www.ascilite.org.au/conferences/melbourne08/procs/rourke.pdf>.
- Saldaña, J. (2016). *The coding manual for qualitative researchers*, (3rd. ed.,). London: SAGE.
- Sengupta, S. (2001). Exchanging ideas with peers in network-based classrooms: An aid or a pain? *Language Learning and Technology*, 5(1), 103–134.
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147–170. [https://doi.org/10.1016/S1060-3743\(00\)00022-9](https://doi.org/10.1016/S1060-3743(00)00022-9).
- 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>.
- Weigle, S. C. (2002). *Assessing writing*. New York: Cambridge University Press.
- Wu, W. C. V., Petit, E., & Chen, C. H. (2015). EFL writing revision with blind expert and peer review using a CMC open forum. *Computer Assisted Language Learning*, 28(1), 58–80. <https://doi.org/10.1080/09588221.2014.937442>.
- 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.
- 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>.
- Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quarterly*, 16(2), 195–209. <https://doi.org/10.2307/3586792>.

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
