Different dimensions of teacher and peer assessment of EFL learners' writing: descriptive and narrative genres in focus

Natasha Pourdana[*] and Soheila Asghari

* Correspondence: Natasha.qale@gmail.com
Department of ELT and English Translation, Faculty of Literature and Foreign Languages, Islamic Azad University, Karaj Branch, Moazen Blvd., P.O. Box 31485-313, Karaj, Iran

Abstract

The present study explored the extent to which teacher assessment (TA) and peer assessment (PA) differ in terms of magnitude and patterns of distribution across English as a Foreign Language (EFL) learners' descriptive and narrative writing performance. Twenty Persian-speaking EFL learners were non-randomly selected and voluntarily participated in a 12-session writing course at a private language school in Iran. Their performance on descriptive and narrative writing tasks was subjected to PA and TA sequentially. The West Virginia Department of Education descriptive writing rubric (with five components of organization, development, sentence structure, word choice and grammar, and mechanics) and Smarter Balanced narrative writing rubric (with five categories of narrative focus, organization, elaboration of narrative, language and vocabulary, and conventions) were adopted to schematize and analyze the distribution of the TA and PA comments. The results of frequency analysis indicated that TA far outnumbered PA on both descriptive and narrative genres of writing. Furthermore, on both descriptive and narrative writings, TA and PA commentaries were local in scope, form-focused, fluctuating, and inconsistent. Also, the distribution of TA and PA comments mainly focused on the conventions of narrative writings, while other macro-components of narrative writing had a steady and depleted pattern. The statistical results confirmed the significance of the observed differences between the number and the nature of TA and PA on descriptive and narrative genres of writing. The researchers made their concluding remarks on the probable causes of observed diversities, imposed limitations of the study, and a number of topics for future research.

Keywords: Descriptive writing, Genre-based, Narrative writing, Peer assessment, Teacher assessment

Introduction

Classroom assessment has been known as an immediate means to attest student learning and progress (Zhao & Liao, 2021). Over the last decades, a paradigm shift was seen in teaching practitioners' tendency, from using assessment to grant credits and grades to students (Liu & Huang, 2020) to a more integrated endeavor of investment on students optimized learning (Zhao & Zhao, 2020). Classroom writing assessment in English as a

Foreign Language (EFL) contexts, however, is still dominated by language teachers' use of writing tasks to make summative judgments about student performance (Wang, Lee, & Park, 2020). Given the potential of assessment to "scaffold" the students writing, it seems that assessment should be re-conceptualized as the central pillar of classroom writing.

At the heart of assessment for learning (AfL) (William, 2018), teacher assessment (TA) and peer assessment (PA) are both known as instructional methodologies and "dialogic learner-centered" assessment tools of writing (Carless et al., 2011; Tian & Zhou, 2020; Yu, 2020). Calling for further research, previous studies on EFL writing have demonstrated controversies over the alignment of TA and PA from different perspectives, such as the plainness of language in peer feedback relative to teacher corrective feedback (Mao & Crosthwaite, 2019), relative success of PA over TA in raising L2 writers' evaluative judgment (Min, 2016), or "the capability to make decision[s] about the quality of work of self and others" (Tai, Ajjawi, Boud, Dawson, & Panadero, 2018, p. 471) and substantial role of PA in the formation of scholarly community of students in their genre-based writing courses (Yu, 2020). On the other hand, several studies on EFL writing argued the efficacy of TA from an achievement angle (Yu & Lee, 2016), where students progress was purely evaluated through analysis of their pre- and post-test gain scores. They encouraged training students to active peer feedback (Mohamadi, 2018) and demanded future research on alternative assessment modalities.

Several studies on genre-based writing assessment in EFL/ESL contexts favored the strict and formal teacher corrective feedback and undermined PA practice for its incompatibility to the exam-oriented educational systems, due to its time-consuming and yet less effective nature (Birjandi & Hadidi Tamjid, 2012; Zhao, 2018). Embedded as a complement to TA routine in writing courses, PA has been understudied as an effective and collaborative learning and assessment opportunity in L2 classroom setting (Dressler, Chu, Crossman, & Hilman, 2019; Pourdana, Nour, & Yousefi, 2021) which highly requires further examination. In the same vein, with the emergence of genre-based assessment of writing in L2 context and its challenges and controversies, several studies examined how L2 learners performance on different writing genres might be affected by such contextual and individual factors as L2 learners' limited language proficiency (Lee, 2017), concerns over hurting their peer's 'face' (Yu & Lee, 2016), reachability of the target audience (Sari, 2019), or academic context of writing (Hyland, 2003); yet what has left uncovered and under-researched is how genre-based assessment of L2 writing could moderate such reported effects (Yu & Lee, 2016). Last but not the least, research literature on assessing writing securitized it from either linguistic/form-focused perspective, confirming the positive role of writing in improving such qualities in language performance as accuracy (Liao, 2016), fluency (Plakans, Gebril, & Bilki, 2019), and complexity (Lahuerta, 2018; Zenouzagh, 2020), or from psycholinguistic/meaning-focused perspective, supporting the influential role of writing assessment in incitement of such affective factors as anxiety (Fathi & Khodabakhsh, 2020) or demotivation (Yu, Jiang, & Zhou, 2020). Hence, little attention has been paid to genre-based assessment approach to L2 writing performance.

Missing in the research studies is also the expansion of the sociocultural theory (SCT) (Vygotsky, 1980) into a multi-perspective approach to writing assessment, which emphasizes both the assessment process and product, while pays special attention to who (teacher, peer and/or self) would participate in the assessment practice (Ataie-Tabar, Zareian, Amirian, & Adel, 2019). Although the research studies frequently

reported the merits of genre-based assessment approach to developing L2 learner writing achievement (Chen & Su, 2012), only a few studies examined the nature and magnitude of teacher and peer commentaries on genre-based writing in L2 environment (Hyatt, 2005; Mirador, 2000; Yelland, 2011).

This study intended to fill the void by adopting a learner-oriented sociocultural approach to genre-based assessment of writing in EFL context, investigating the parallel potentials of TA and PA on descriptive and narrative writing genres. The researchers' goal for choosing descriptive genre of writing was to examine the EFL learners' ability in describing facts, figures, and flowcharts in written mode. Likewise, the researchers' goal for choosing narrative genre of writing was to monitor EFL leaners' writing performance on recounting facts, and reporting step-wise processes. In the following section, a critical review of the studies on teacher and peer feedback to assess writing is presented. Next, the objectives of the study are restated, followed by the procedure of data collection. Then, the statistical results are discussed and consolidated with previous research findings. Finally, drawn conclusion, implications and suggestions for further research are presented in light of limitations of the study.

Literature review

Teacher versus peer assessment

As a core component in every EFL program, teaching-to-write has intrigued many L2 teachers and practitioners. In the same vein, learning-to-write is a necessity which boosts student access to teacher assessment and its constructive impacts. Therefore, on occasions where giving feedback to students is not part of language teachers' formal practice or when language learners fail to intake the received feedback on their writing, reaching the intended goals of writing in EFL seems an impossibility (Agbayahoun, 2016).

A considerable body of research in EFL writing has documented the importance of teacher assessment (TA). As a pedagogical genre, TA is manipulated to transfer a heavy informational load to the students, to provide comments on the form and content of a written text and to encourage students to develop their writing (Karimi & Asadnia, 2015; Sermsook, Liamnimitr, & Pochakorn, 2017). The crucial role of TA is commonly acknowledged in language classroom where full access to target language is scarce and where a product-oriented approach towards assessing writing is formally adopted. Agbayahoun (2016)'s research on EFL student writing in the Republic of Benin indicated that although TA promoted students' language accuracy and mechanics of writing, they favored teacher comments more on "the content" of their written products. Similarly, in another experimental study by Luquin and Garcia Mayo (2021) on teacher written feedback in terms of providing written models—the native-like texts to which young learners can compare their own writing—the researchers analyzed children's pair talk for evidence of any content and linguistic problems. It was found that the students looked more for content-related than linguistic features at the comparison stage. Moreover, these children actively incorporated teacher comments on mechanics and discourse-related features into their rewriting stage (Luquin & Garcia Mayo, 2021), relative to the control group who self-revised their written scripts.

As an alternative assessment genre, peer assessment (PA) has received special attention by education researchers over the past decades. PA—also known as peer evaluation, peer

feedback, peer review, or peer critique—refers to "a communication process through which learners enter a dialog related to their own and peer performance" (Liu & Carless, 2006, p. 280). Despite ample theoretical and empirical support, the effectiveness of peer feedback in various learning environments has remained disputable (Chang, Lee, Tang, & Hwang, 2021). In teacher-centered educational systems, where students often display low ability in self-study, in autonomous problem-solving, and in adapting effective learning strategies, PA can be their remedy (Deng & Carless, 2010). Collaborating in PA, students can generate opportunities to discuss their written texts (Yu & Lee, 2016), improve the quality of their writing (Zhao, 2014), promote their motivation to write (Shih, 2011), and enhance their critical thinking (Joordens, Pare, & Pruesse, 2009).

Although numerous studies reported the educational benefits of PA, some research results suggested that it might hardly be useful as formal assessment in real language classrooms (Brown, 2004; Li, 2017; Liu & Li, 2014; Pope, 2001). In fact, there is little agreement over the adoptability of PA as an alternative assessment to TA. Chang et al. (2021)'s meta-analysis of 20 peer assessment articles and their 340 follow-up studies reported the heterogeneous effect size values for efficacy of PA. In one semester-long intervention study of the efficacy of PA on improving English writing of 70 sophomore Chinese students, and promoting their learning autonomy, Shen, Bai, and Xue (2020) concluded that PA enhanced the students' learner autonomy in a considerable way. They believed that learner dependency on the teacher feedback was noticeably reduced and the students' confidence in their learning ability was boosted as a result of their active PA practice. Several research findings reported significant contribution of PA in L2 contexts, in terms of better understanding of the learning objectives (Hattie & Timperley, 2007), regulating learners own performance (Colognesi & Deschepper, 2018), and increasing student metacognitive self-awareness (Lu & Law, 2011). Matsuno (2009) also suggested that PA practice remains challenging and engaging in L2 learning contexts and would cause constructive interactions among language learners.

From a statistical perspective, Cheng and Warren (1999) reported a strong interrelation between teacher raters and peer raters, depending on the target tasks and the context of language use (Jafarpur, 1991; Patri, 2002; Saito & Fujita, 2004). They concluded that proper training could improve the quality and magnitude of PA commentaries in a substantial way (Saito, 2008). Moreover, Jones and Alcock (2014) obtained a high validity and inter-rater reliability measures for PA practice, asking students to compare their written scripts with their peers, and concluded that the students could easily become effective peer assessors. On the other hand, in a comparative study of TA and PA with 38 EFL university students academic writing, Azarnoosh (2013) implemented a learner attitude questionnaire and a friendship bias analysis with a pretest/posttest writing research design. She observed no significant differences between the impacts of PA and TA on student academic writing. Likewise, no friendship bias was detected in student PA comments, yet a positive change in students' attitudes towards PA emerged.

Descriptive and narrative genres of writing

The concept of genre-based assessment of L2 writing in terms of student performance on writing life narrative or describing a procedure in a research report has remained under-researched (Jaubert, 2007). A number of studies have focused on genre-related

teacher assessment of L2 writing performance in colleges and universities (Dolz & Gagnon, 2008; Freedman & Medway, 1994; Horverak, 2016; Hyland, 2003, 2007). In addition, the formative assessment of genre-based writing, such as expository, narrative, and descriptive essays, was reported as a constant challenge to teacher assessment (Dumais, Lafontaine, & Pharand, 2017). Teachers often "have difficulty, resistance in carrying out certified, and even formative evaluation of pupils' language production" (Gagnon, De Pietro, & Fisher, 2017, p. 16).

By definition, "genre" not only refers to different types of literary texts, such as poetry or drama, but also to "the conventional and recurring patterns of everyday, academic and literary texts within a particular culture" (Derewianka, 2003, p. 133). Members of a particular social and cultural community are expected to recognize the generic language conventions in various situations and be able to respond by using the appropriate genre. Spoken or written, genres are often identified according to five primary social purposes of recount or description, narrative, information report, instruction, explanation, and expository (Swales, 1990).

In general, descriptive genre gives detailed attributes to a person, place, or event. Those entities should be described in such a way that the reader can picture the topic and enter the writer's experience. Descriptive genre is also considered as a way to improve other genres of writing, such as analytical and argumentative or perhaps as a dominant strategy to depict an image of what something looks like (Birjandi & Hadidi Tamjid, 2012). In descriptive writing, the writer introduces the intended theme in the topic sentence before describing it further through the supporting sentences. Topic sentence often contains "enumeration," and the use of enumerators notifies what pieces of information is being presented in the supporting sentences to the readers. Characteristics, parts, aspects, layers, formats, habits, behaviors, and facets are the most frequently used enumerators in descriptive genre of writing.

One of the well-known rubrics for assessing descriptive writing has been developed by West Virginia Department of Education (WVDE) (2011). WVDE rubric consists of five categories of organization, development, sentence structure, word choice and grammar, and mechanics. "Organization" entails a clear and logical progression of ideas in a descriptive writing. "Development" is the clear focus maintained by the writer for the intended audience, by the strong use of examples, relevant details, analogies, and illustrations. "Sentence structure" refers to the use of well-constructed sentences with varied structures. "Word choice and grammar" is the writer's preference for vivid words and grammatical phrases. Finally, "mechanics" refers to the systematic use of punctuation, capitalization, and grammar in writing (NBCT Office of Assessment West Virginia Department of Education, 2015).

Narrative genre, as the art or gift of storytelling, is made through every moment of our lives so that we make narration plenty of times every day (Abbott, 2002). Lou-Conlin (1998) defined written narrative as a gradual mode of development through which the writer expresses the chronological sequence of ideas and events. Narration is mostly done with the purpose of maintaining the readers' interest in the course of actions to depict a given event or personal experience narrative (PEN) (Labov, 1997). To fulfill this purpose, the narrator needs to illustrate the timeline of the experience with some extraordinary meaning or significance to both the reader and the writer.

Developed by Tulare County Office of Education in 2014, Smarter Balanced narrative writing rubric has been widely used for assessment of written narratives. The rubric identifies five components of narrative focus, organization, elaboration of narrative, language and vocabulary and conventions. "Narrative focus" refers to the effective establishment of a setting, narrator and/or characters. "Organization" entails creating an effective plot which demands unity and completeness. "Elaboration of narrative" is the thorough and effective elaboration using details and dialogs. "Language and vocabulary" means deliberate choice of words and structures that expresses experiences or events. Finally, "convention" indicates the effective and consistent use of punctuation, capitalization, and spelling (Smarter Balanced Assessment Consortium, 2012).

To echo the importance of genre-based writing assessment in EFL contexts, a number of studies documented how writer performance has been modified by corrective feedback on various writing genres, such as argumentation, narration and analysis. Zareei (2009) studied the impact of the genre-based TA on the quality and quantity of 140 Asian EFL learners writing of the letters of application. She reported that while the writing accuracy and complexity were significantly improved, the quantity (i.e., fluency) in students writing performance remained unaffected. Likewise, Rezvani, Aqdam, and Saeidi (2014) supported the positive role of teacher genre-based formative assessment on the achievement of advanced EFL learners in an experimental pretest/posttest research study. In another study, the impact of the genre-based TA on Colombian EFL students' argumentative writing was examined through questionnaires, semi-structured interviews, class recordings, and students' artifacts which indicated a significant enhancement in student self-confidence and their highly positive attitudes towards writing (Chala Bejarano & Chapetón, 2013). Genre-based assessment was also reported to provide "scaffolding" opportunities through dialogic interactions by the students which further supported students understanding of writing as a process (Yu, 2020).

A number of recent studies have explored the role of genre-based peer feedback in L2 writing development by comparing it to teacher corrective feedback, in reducing treatable language errors of EFL learners (Diab, 2015), and L2 learners' perceived importance of peer feedback to their academic writing development (Seror, 2011). In another study, Yang et al. (2006) reported that the EFL learners used 90% of teacher corrective commentaries in revising their expository essays, while they incorporated only 67% of peer reviews. They commented on TA as "very useful" and "informative," while being critical of PA as "inaccurate," "confusing," or "judgmental." Xu and Liu (2010) compared the nature of peer and teacher assessment of EFL student writing and concluded that PA mostly focused on micro-level language features, while TA addressed the form, content, and mechanics of writing. Overall, research reports suggested that both TA and TA would have significant roles to play in L2 learner writing improvement. However, few studies elaborated on the nature of observed differences between PA and TA comments on the EFL genre-based writing performance.

This study

To bridge the gap in the research literature, this study was intended to explore two dimensions of the teacher and peer assessment, in terms of the relative quantity of the TA and PA comments, and the nature of differences in TA and PA comments on EFL

student descriptive and narrative writing performance. To fulfill the objectives of the study, three research questions were raised, as following:

(1) How teacher assessment and peer assessment differ in their impacts on EFL learners' descriptive writing?
(2) How teacher assessment and peer assessment differ in their impacts on EFL learners' narrative writing?
(3) Are the differences between teacher assessment and peer assessment impacts on EFL learners' descriptive and narrative writing significant?

Method

Participants

A sample of 20 EFL learners (15 male and 5 female) who enrolled at a private language institute in Karaj, Iran, took part voluntarily in this study. A non-random purposive sampling method was used in this study (Ames et al., 2019), as selecting a representative group of EFL learners with adequate experience in genre-based writing in English was the researchers' plan. Normally, there is no systematic emphasis on genre-based writing in curricular programs at private language institutes in Iran. However, the participants in this study were already required to submit at least 10 writing tasks, such as writing an invitation card, a letter of application or diaries in previous courses as partial requirement of their text book *Top Notch 3* (Saslow & Ascher, 2006). Their experience in learning English was between 3 to 6 years ($m = 5$) and their ages ranged from 17 to 27 ($m = 20.25$). The participants' language proficiency level at the moment of conducting this research was intermediate (30-37, B1 in OPT ranking system). Additionally, two MA graduates of English language teaching (ELT) took part in this research as the EFL teacher and the co-raters of the participants' descriptive and narrative writing. Both had at least 7 years of teaching experience at private language schools in Iran. Table 1 summarizes the demographic information of the EFL learner participants.

Instruments

Oxford placement test (OPT)

Developed by Oxford University Press and University of Cambridge Local Examinations Syndicate (CLES), OPT (Version 1, 2001) has been calibrated against the English proficiency levels provided by Common European Framework of Reference (CEFR) on a six-level assessment scale, ranged from A1 (lowest level) to C2 (highest level). In this study, OPT was used as a placement test aiming at measuring the selected participants' knowledge of general English. The participants were required to answer the 60 multiple-

Table 1 EFL learners' demographic information

Participant	Gender	Age range	Major	Studying English	OPT score range
n = 20	Female, (5) 25%	17–20, (16) 80% 21–24, (2) 10% 25–27, (2) 10%	High school students (10) 50% College students (6) 30% University students (4) 20%	3–4 years (12) 60% 5–6 years (8) 40%	30–33 (11) 55% 34–37 (9) 45%

choice items in the test in 45 min. Their language proficiency level was determined as intermediate (30-37, B1 in OPT ranking system) (Cronbach's α = .812).

West Virginia Department of Education (WVDE) descriptive writing rubric

In order to assess the participants' descriptive writing in both TA and PA practice, the well-known West Virginia Department of Education rubric (2011) for descriptive writing was adopted. The reason behind adopting the WVDE descriptive writing rubric was its user-friendliness and creditability. WVDE rubric is well-known for meeting the criteria to assess English descriptive writing, by defining rigorous and clear-cut score bands to ensure an accurate representation of student genre-based writing performance (NBCT Office of Assessment West Virginia Department of Education, 2015). WVDE descriptive writing rubric consists of five categories of "organization, development, sentence structure, word choice and grammar, and mechanics," within 1–6 band scores, ranging from 1 (minimal) to 6 (exemplary) spectrum (Additional file 1: Appendix A)

Smarter balanced narrative writing rubric

Known as a public agency, the Smarter Balanced Assessment Consortium (SBAC) has created a global digital library of formative assessment tools to support teachers and students. Developed by Tulare County Office of Education (CA, USA) in 2014, Smarter Balanced narrative writing rubric has since been used to assess language learners' written narrative performance. The Smarter Balanced assessment system includes a comprehensive suite of standard summative and formative assessment tools that save writing teachers time (Smarter Balanced Assessment Consortium, 2012). The logic behind selecting this rubric in this study was its clarity and user-friendliness which made it more accessible to peer assessment. Smarter Balanced narrative writing rubric consists of five categories of "narrative focus, organization, elaboration of narrative, language and vocabulary and conventions." The rubric has a 1 to 5 band scores, ranging from 0 (no evidence of the ability to write a narrative) to 5 (meeting all the criteria of writing a real or imagined narrative) (Additional file 1: Appendix B).

Descriptive and narrative writing elicitation tasks

The designed interventions in this study lasted for 12 sessions in six consecutive weeks. The participants received six descriptive (on the odd sessions) and six narrative (on the even sessions) writing elicitation tasks. The arrangement and topics of writing tasks are summarized in Table 2.

Data collection procedure

Prior to the experiment, all the participants were asked to sign a consent form for their volunteer participation in the study. Data collection procedure was conducted in the following steps:

- Two days before the writing course began, an OPT Version 1 (2001) was administered as the placement test for the purpose of normalizing the participants for their language proficiency level. Counting on the OPT scores, 20 participants

Table 2 Topics and arrangement of the descriptive and narrative writing elicitation tasks

Session	Genre	Writing task
1	Descriptive	What are your reading habits?
2	Narrative	What was your best vacation?
3	Descriptive	How do you play with your smartphone?
4	Narrative	What was the last time you saw a doctor?
5	Descriptive	What type of movies do you like most?
6	Narrative	What do you remember from your childhood?
7	Descriptive	How is your neighborhood?
8	Narrative	What do you remember from the last birthday party you were invited?
9	Descriptive	What is your ideal apartment?
10	Narrative	How does your life look like in 5 years?
11	Descriptive	How do you clean your room?
12	Narrative	What do you remember about the worst day of your life?

were selected out of 31 whose scores were within the band scores of 30–37 (B1, in OPT ranking scale).

- A day before the experiment, the researchers provided the participants with an 8-h tutorial on the elements of descriptive and narrative genres of writing as well as the two selected rubrics for genre-based writing assessment. The participants were instructed on the nature of peer feedback in terms of selective comments, evaluation, or suggestions they could provide on each other's writing. They were informed that PA was not substituted for TA, and they would receive teacher comments on their writing every session.

- A topic familiarity checklist was prepared by the researchers and distributed among the participants, a day before the experiment. The purpose was to select and incorporate 12 most favorable topics into descriptive and narrative writing elicitation tasks.

- The 6-week writing course was divided in to six odd sessions devoted to descriptive writing tasks and six even sessions on narrative writing tasks. The participants were assigned into pairs who were required to individually write a 200 to 250-word paragraph on the assigned topics, followed by reading peer's written draft and giving feedback. They were allowed to consult the WVDE descriptive writing rubric, and Smarter Balanced narrative writing rubric, in the PA process. Drafting and PA lasted for 60 min.

- At the end of every session, the teacher collected the writings for her formative assessment. The participants were asked to revise/re-write their first draft according to the received PA and TA before submitting the final draft in the following session. As an in-class post-writing task, the students received no grade or bonus for their PA practice.

- Teacher assessment was carried out collaboratively by the two raters in the study (see Additional file 1: Appendices C and D for selected samples of descriptive and narrative writing, along with TA and PA comments). The possible controversies in rating written drafts were discussed and resolved upon every occasion. The inter-rater reliability indices were calculated for a variety of ratings in this study (Cronbach's α = 0.881, reliability index for descriptive writing tasks; Cronbach's α =

0.781, reliability index for narrative writing tasks; Cronbach's α = 0.681, reliability index for TA and PA ratings of descriptive writing tasks; Cronbach's α = 0.901, reliability index for TA and PA of narrative writing tasks). All values of Cronbach's alpha represented strong agreement and statistical significant ($p < .05$).

Coding system

As indicated in Table 3, the TA and PA comments on student descriptive writing were cracked down into 11 themes with reference to five categories in WVDE descriptive writing rubric. Along the retrieved themes, examples of TA and PA comments were given in Table 3.

Similarly, TA and PA commentaries on narrative writing tasks were schematized based on the five categories in Smarter Balanced narrative writing rubric. They were further decoded into 11 themes in Table 4. The themes and examples of TA and PA comments were given in Table 4.

Results

The IBM SPSS 23 was employed for statistical analysis of the collected data. Every comment, correction, or suggestion made by the teacher or peers on the written drafts was considered as one "feedback point" and the total feedback points for every draft of writing were calculated. Therefore, the observed *decrease* in teacher and peer feedback points was interpreted as the student progress in their writing. A descriptive statistical analysis was conducted with the PA and PA counts on participants' descriptive writings, produced on the six odd sessions. As it can be seen in Table 5, the mean scores were much higher in TA than in PA, ranging from m = 12.80 on session 1 to m = 24.40 on session 9. On the other hand, PA had its minimum mean score on session 3 (m = 5.40) and the highest mean score on session 9 (m = 13.00). Regarding the measures of standard deviation, PA enjoyed smaller standard deviations (from SD = .96 on session 1 to SD = 4.49 on session 9) than the TA (from SD = 1.90 on session 1 to SD =

Table 3 The analytical scheme for classifying TA and PA on descriptive writing tasks

Category	Theme	Teacher feedback Example	Peer feedback Example
Organization	• Transitional devices	"The sentences need a connector, therefore?"	"I think the sentences are very short."
Development	• Details	"Add adverbs to modify the verb **work.**"	"……. Working hard"
Sentence structure	• Compound sentence • Complex sentence • Subject-verb agreement	"Dangling subject!" Who is the doer in this sentence?	"They ……… have cooperated."
Word choice and grammar	• Singular vs plural nouns • Adjective vs adverb • Pronoun	"activitiveness? Check your dictionary."	"The tenses are not correct."
Mechanics	• Capitalization • Punctuation • Spelling	"however; it needs semicolon."	"Names of the seas are capital."

Table 4 The analytical scheme for classifying TA and PA on narrative writing tasks

Category	Theme	Teacher feedback Example	Peer feedback Example
Narrative focus	• Time • Place • Character	*"It is confusing! Where and when?*	*"When did he see the teacher?"*
Organization	• Transitional devices • Sentence formation	*"It is called enumerators. You should use them."*	*"He saw her first."*
Elaboration of narrative	• Details • Sensory words	*"It was fun? Explain more."*	*"Use adjectives like larger, bigger…."*
Language and Vocabulary	• Figurative language	*"Taller or higher?"*	*"The past tense of lie is laid?"*
Conventions	• Capitalization • Punctuation • Spelling	*"Read about the use of comma in conditional sentences."*	*"Proper nouns should be capital."*

5.93 on session 5) which signaled less variance in PA comments on descriptive writings.

Figure 1a, b displays the distribution patterns of TA and PA comments on participants' descriptive writing, following five components in WVDE descriptive writing rubric. As it can be seen in Fig. 1, both TA and PA focused on the component of "mechanics" of writing at the expense of other components of "organization" or "development." In both TA and PA cases, a fluctuating pattern of comments can be seen along the treatment sessions, with a decrease of the feedback points on all components of descriptive writing by final session of the treatment.

Figure 2 illustrates the observed difference between the total TA and PA feedback points on student descriptive writing. As it can be seen, along the six sessions of treatment, the sum of TA feedback points has been growing up from $m = 12.80$ on session 1 to $m = 24.40$ on session 9, following a pullback on the last session ($m = 15.20$). On the other hand, with more fluctuation and irregularities, PA feedback points began to rise from $m = 7.80$ on session 1 to $m = 13.00$ on session 9 and a noticeable regression to $m = 6.00$ on session 11.

TA and PA feedback points on participants' descriptive writing showed similarities and differences in their magnitude and patterns of distributions. The quantity of the feedback points in TA exceeded PA on all five components of descriptive writing in WVDE rubric, throughout the six sessions of treatment. Moreover, both TA and PA feedback points followed similar fluctuating pathways, paying laser-focus to the form-focused component of "mechanics of writing" and marginal attention to other four global components such as "organization" or "development."

Another descriptive statistics was carried out with TA and PA feedback points on student narrative writing, produced and collected on the six even sessions. As it can be
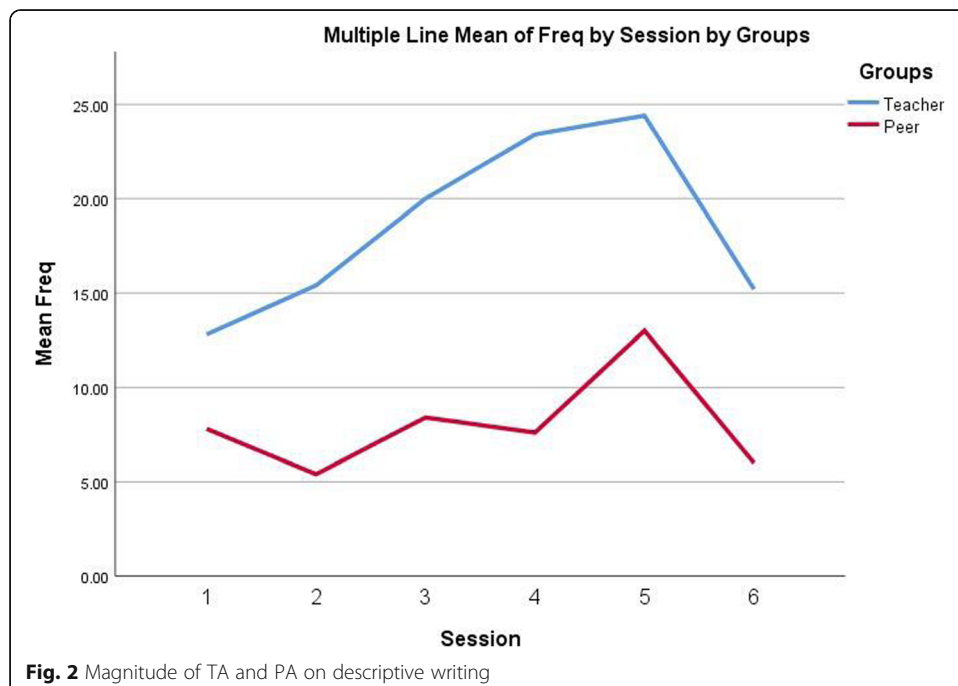
Table 5 Descriptive statistics of TA and PA on descriptive writings

Assessment type	Session 1		Session 3		Session 5		Session 7		Session 9		Session 11	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
TA	12.80	1.90	15.40	3.70	20.00	5.93	23.40	3.58	24.40	6.36	15.20	5.17
PA	7.80	.96	5.40	1.69	8.40	3.31	7.60	2.67	13.00	4.49	6.00	2.32

Fig. 1 TA (**a**) and PA (**b**) patterns of distribution on descriptive writings
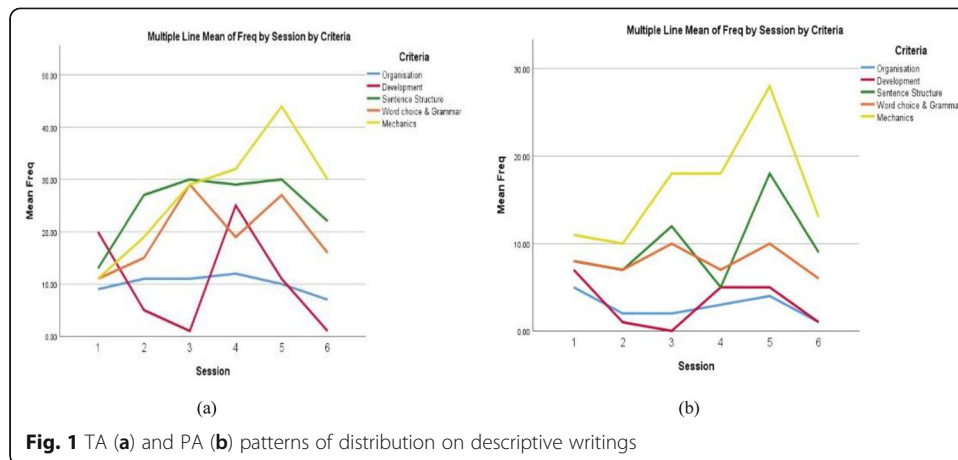
seen in Table 6, the mean scores were higher in TA than in PA, ranging from m = 11.80 on session 8 to m = 24.80 on session 10. PA also had its smallest mean score (m = 7.20) on session 12 and the highest mean score (m = 15.40) on session 10. Once again, PA enjoyed smaller measures of standard deviation (from SD = 5.06 on session 8 to SD = 10.92 on session 10) than TA (from SD = 5.84 on session 8 to SD = 14.37 on session 10), which suggested less variance in PA comments on narrative writings.

Figure 3a, b reexamines the distribution of TA and PA feedback points on students narrative writing, following five components in Smarter Balanced narrative writing rubric. Both TA and PA feedback points were divided into two areas, so that the majority of TA and PA comments escalated the "conventions" of narrative writing, while minimum attention was paid to other macro-components such as "narrative focus" or "organization" of narrative. While both TA and PA comments showed a noticeable

Fig. 2 Magnitude of TA and PA on descriptive writing

Table 6 Descriptive statistics of TA and PA on narrative writings

Assessment type	Session 2		Session 4		Session 6		Session 8		Session 10		Session 12	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Teacher	15.00	8.36	13.00	7.59	13.80	7.10	11.80	5.84	24.80	14.37	16.80	9.42
Peer	8.60	6.60	7.00	5.83	9.60	7.15	5.80	5.06	15.40	10.95	7.20	5.73

pattern of rise and fall on "narrative convention," the feedback points on other components of narrative writing remained marginal and steady from the beginning to the end of treatment sessions.
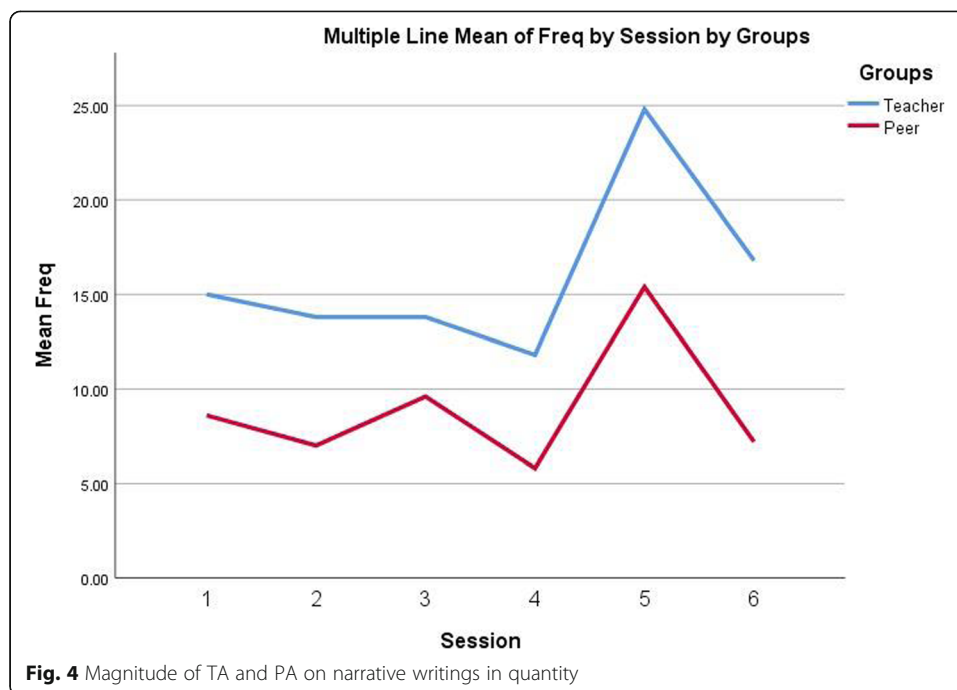
Figure 4 illustrates the observed difference between the total TA and PA feedback points on student narrative writing. The patterns of TA and PA distribution were less divergent and more proportional on student narrative writing. Through the six sessions of treatment, the quantity of TA feedback points has been sinking from $m = 15.00$ on session 2 to $m = 11.80$ on session 10, rising up to $m = 14.37$ on session 10 and falling to $m = 16.18$ on the last session. With less irregularities, PA began from $m = 8.60$ on session 2, sliding down to $m = 5.80$ on session 8, ebbing to $m = 15.40$ on session 10, and flowing to $m = 7.20$ on session 12.

The distribution of TA and PA feedback points on participants' narrative writing showed similarities than differences. The feedback points of TA outnumbered the PA on the five components of narrative writing, on all six sessions of treatment. Meanwhile, both TA and PA moved along a similar snaking pathway, paying their highest attention to "convention" of narrative writing, while equally focused less on the other global components of written narrative.

To compare the observed differences in TA and PA feedback points on descriptive and narrative writings, the researchers conducted the cross-tabulation of TA and PA feedback points (i.e., group) on the components (i.e., selected criteria), indicated in WVDE descriptive writing rubric and Smarter Balanced narrative writing rubric (Table 7). It was followed by a Pearson chi-square test to determine the statistical significance of the present divergence in the distribution of TA and PA feedback points on descriptive and narrative writings (Table 8).

Table 7 is horizontally divided into components of descriptive writing (upper) and narrative writing (lower), while it is vertically partitioned into TA and PA feedback points and percentages in the two right columns. The overall magnitude of the TA and

Fig. 3 TA (**a**) and PA (**b**) patterns of distribution on narrative writings

Fig. 4 Magnitude of TA and PA on narrative writings in quantity
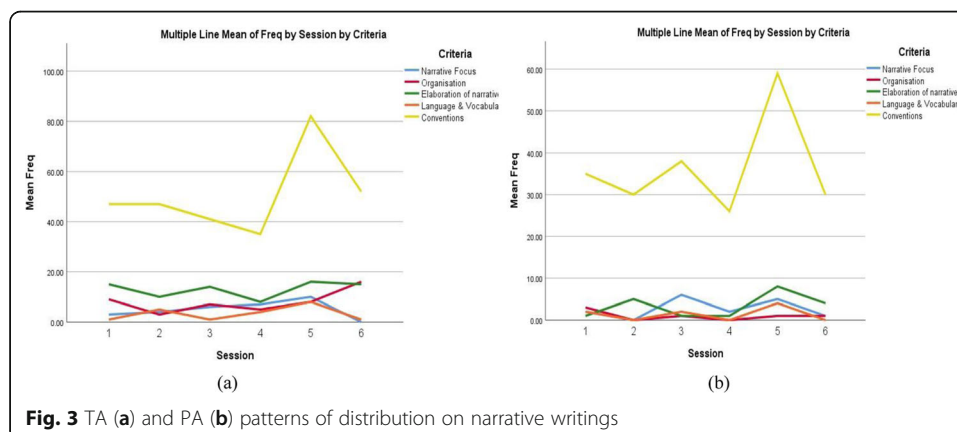
PA feedback points for both descriptive and narrative writing indicated remarkable convergence:

– On the student descriptive writing, both TA and PA feedback points were maximum on "mechanics of writing" ($n = 158$ and $n = 80$, respectively) and minimum on "organization" ($n = 53$ and $n = 17$, respectively).
– On student narrative writing, both TA and PA total feedback points were maximum on "conventions of narrative writing" ($n = 267$ and $n = 190$, respectively) and minimum on "elaboration of narrative" ($n = 16$ and $n = 9$, respectively).

Table 7 also summarized the total number of TA feedback points ($n = 968$, 68%) which was more than twice the size of PA feedback points ($n = 440$, 31%). The proportions of TA and PA feedback points across descriptive and narrative writings were statistically probed with a Pearson chi-square.

As illustrated in Table 8, the value of Pearson chi-square was statistically significant ($X^2 = 50.000$; $df = 19$; $p = .000 < .05$, Cramer's $V = .331$), which indicated the contingency of the results. In other words, the observed differences between TA and PA comments were meaningful and statistically significant on both descriptive and narrative writing genres.

Discussion

Language assessment plays a critical role in facilitating students' learning process and learning outcomes. The new paradigm of assessment for learning (AfL) conceptualizes alternative assessment as a teacher-student or student-student "dialogic process." To observe the differences in the quality and quantity of teacher assessment (TA) and peer

Table 7 Cross-tabulation of components in descriptive/narrative writings x TA/PA

Criteria			Group	
			Teacher assessment	**Peer assessment**
Descriptive writing	Organization	Count	53	14
		% within criteria	79.0%	20.0%
		Std. residual	1.0	− 1.0
		Adjusted residual	1.0	− 1.0
	Development	Count	57	17
		% within criteria	80.0%	19.0%
		Std. residual	1.0	− 1.0
		Adjusted residual	2.0	− 2.0
	Sentence structure	Count	156	54
		% within criteria	74.0%	25.0%
		Std. residual	.0	− 1.0
		Adjusted residual	1.0	− 1.0
	Word choice and grammar	Count	125	41
		% within criteria	75.0%	24.0%
		Std. residual	1.0	− 1.0
		Adjusted residual	1.0	− 1.0
	Mechanics	Count	158	80
		% within criteria	66.0%	33.0%
		Std. residual	.0	.0
		Adjusted residual	.0	.0
Narrative writing	Narrative focus	Count	23	14
		% within criteria	62.0%	37.0%
		Std. residual	.0	.0
		Adjusted residual	.0	.0
	Organization	Count	70	19
		% within criteria	78.0%	21.0%
		Std. residual	1.0	− 1.0
		Adjusted residual	2.1	− 2.1
	Elaboration of narrative	Count	16	8
		% within criteria	66.0%	33.0%
		Std. residual	.0	.0
		Adjusted residual	.0	.0
	Language and vocabulary	Count	43	16
		% within criteria	87.0%	12.0%
		Std. residual	1.0	− 2.0
		Adjusted residual	2.0	− 2.0
	Conventions	Count	267	190
		% within criteria	58.0%	41.0%
		Std. residual	− 2.0	3.0
		Adjusted residual	− 5.0	5.0
Total		Count	968	440
		% within criteria	68.0%	31.0%

Table 8 Pearson chi-square test of TA and PA x descriptive and narrative writing

Chi-square test			
	Value	df	Asymp. Sig. (2-sided)
Pearson chi-square	50.000	19	.000
Likelihood ratio	52.023	19	.000
Linear-by-linear association	23.000	1	.000
Cramer's V	0.331		.005
N of valid cases	1408		

assessment (PA), the researchers in this study aimed at investigating a genre-based assessment of EFL learners' descriptive and narrative writing. The discussion of the first and second objectives of the study which separately explored the possible differences between TA and PA on the EFL learners' descriptive and narrative writing can be divided into two aspects of the magnitude of TA and PA and their pattern of distribution for five components in WVDE descriptive writing rubric and five components in Smarter Balanced narrative writing rubric. Statistical frequency analysis showed that while the average TA was twice more than PA in quantity, both TA and PA comments were distributed unevenly and unsystematically along the 12 treatment sessions. Moreover, it was found that the majority of TA and PA comments were microscopic, form-focused, and more targeted "the mechanics or conventions of writing" than any other macroscopic and content-based writing criteria, such as "organization or narrative focus."

The findings implied that despite the predictable insufficient skills and knowledge of genre-based writing, the students could engage in assessing their peers' work and critically evaluate the quality of their writing from a descriptive and narrative-genre perspective. Moreover, the similarities between TA and PA distribution suggested that the participants' peer feedback was based on a constant reference to the multiple components in the selected writing rubrics and their close observation of the received teacher feedback. The findings were consistent with Man, Xu, and O'Toole (2018) and Yu (2020) who reported L2 student systematic access to a variety of resources in preparing their PA, including writing rubrics, peer interaction, teacher's anchor writing, and teacher-generated written feedback. The findings provided further supports for the socio-cultural nature of PA. They also indicated that the students paid unequal attention to the components of genre-based writing and provided more form-focused than content-oriented feedback. This unbalanced and fluctuating nature of PA comments was also noticed by Yang and Carless (2013). They reported the participants' low skills in genre-based writing as the main cause for such irregularities in their peer feedback. The students' lack of agency to assess their peers' language production or to generate evaluative judgment on genre-based writing was also speculated in Yu (2020).

The third objective in this study further probed different distributions of TA and PA comments on descriptive and narrative genres of writing. Significance of the observed statistical interactions was confirmed with (a) the cross-tabulation of the 10 criteria (five categories of descriptive writing, and five categories of narrative writing) in two modes of assessment (TA and PA) and (b) running Pearson chi-square contingency test. In other words, the existing difference in TA and PA nature and number was meaningful and would cause different impacts on students' genre-based writing.

The researchers hypothesized two causes for such discrepancies, including (1) the inefficient teacher's how-to-do tutorial on peer assessment, which made it challenging to the students, and (2) the teacher's limited supervision over the students' inductive ability to infer and apply the criteria of descriptive and narrative writing to their peer feedback. In a similar case study on face-to-face PA, Farahani, Nemati, and Montazer (2019) argued that the "illegitimacy" and relatively low weight that the peer feedback commonly carried in the eyes of the students often made them too critical to cast comments to their peers or too defensive to accept and apply comments from peers. Furthermore, the students' concern over the effects their feedback would have on their peers' feelings and emotions (i.e., the sociocultural aspect of feedback) could be another cause for their underperformance (Yu, 2020). Arguing the simplicity, inaccuracy, and uncertainty of PA comments, Lam (2010) also took a speculative approach to PA and the incorporation of PA comments by the students into their genre-based descriptive or persuasive tasks.

In this study, the components of "mechanics" in descriptive writing and "convention" in narrative writing, which received major focus in TA and PA, entailed the guidelines to preserve the accuracy in punctuation, spelling, and capitalization. Therefore, the current researchers assumed that both TA and PA comments were more microscopic and form-focused than global and content-oriented. The disproportionality of TA and PA comments could optimistically signal the participants' confident performance on the global criteria of genre-based writing, at the cost of a more form-focused and stylistic approach to their production. From a different perspective, Yu (2020) argued that the participants would likely adopt a "passive stance" and have a tendency to focus more on the language form, rather than the genre-based conventions in their written and oral peer feedback. Farahani et al. (2019) also speculated an unwanted synergy or a "sequential skill acquisition" (Schunk & Zimmerman, 1997) between TA and PA, as a consequence of students' close "observation and emulation" of the teacher corrective feedback to increase their evaluative judgments in PA practice. The findings of this study corroborated in to a number of other experimental research as far as the differences between teacher assessment and peer assessment are concerned (Sermsook et al., 2017; Tai, Lin, & Yang, 2015; Wang, 2014; Zarifi, 2017). In addition, the findings were in line with a number of studies that supported the positive roles of both TA and PA in directing learners' attention to different quality aspects of their writing process (Bitchener, 2008; Diab, 2015; Li, Link, & Hegelheimer, 2015; Rassaei, 2013; Yilmaz, 2013).

Conclusion

The findings of this study enriched our understanding of the distribution and disposition of genre-based teacher and peer assessment in EFL context. This study simulated a "multi-perspective" assessment setting in which the solid and outcome-oriented teacher assessment was integrated into dynamic and community-oriented peer assessment in EFL learning environment. The efficacy of each mode of assessment can be predicted with reference to the extent of teacher-student sociocultural interactions, type of language tasks, and the purpose of the course (Farahani et al., 2019). Besides

the critical lead of TA as a standard practice, PA as a shared responsibility will help students adopt a more proactive role to reflect on their own writing and others, to evaluate their performance against a well-defined rubric, to re-draft and revise it, and to learn and improve their own work.

As pointed out by Brammer and Rees (2007), the success of implementing peer feedback heavily relies on how well the assessment criteria and a sense of shared community are built by the teacher in the classroom. Therefore, teachers themselves need systemic training on encouraging PA in a regular and more informative fashion, supplementing their formal and summative assessment. In addition, PA can be more effective when face-to-face peer feedback is accompanied with computer- or mobile-mediated modes of communication (CMC) which enable the language students with synchronous modality of peer feedback. Hence, the language teachers are recommended to explicitly model the procedure of peer assessment, the assessment scale as well as the modem technologies to the students long enough to reassure their acquired skills.

Attempts were made in this study to eliminate the lapses in data collection, content analysis, and data interpretation. However, there are a few limitations imposed on the researchers' drawn conclusions. In this study, 20 EFL learners at the intermediate level of language proficiency were selected with a non-random purposive sampling method. Consequently, the findings in this study should be generalized with adequate cautions in similar context of L2 learning, or with L2 learners at different language proficiency profiles. Still, the findings in this study on how genre-based TA and PA might affect the EFL learners writing can enrich the language teachers' understanding in various educational settings. Secondly, the researchers narrowed their scope down to general TA and PA practices in this study and made no distinction across the types of TA and PA feedback, such as clarification, correction, problem identification, or explanation (Hanjani & Li, 2014) on descriptive and narrative writings. Future researchers are highly recommended to take such strategies into consideration and add more depth and dimensions to their work. Thirdly, the current researchers collected the data exclusively from written TA and PA. Yet, adding the records of face-to-face feedback could log richer data from the students' interactions in pairs or in groups while using (non)verbal modes of communication. Last but not least, before partaking this study, the selected participants had already received formal instructions to genre-based writing in previous EFL courses. Therefore, their writing experience might have compromised their current writing performance to the extent that would affect the generalizability of the results in this study.

Abbreviations
EFL: English as a Foreign Language; OPT: Oxford Placement Test; PA: Peer assessment; TA: Teacher assessment; WVDE: West Virginia Department of Education

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s40468-021-00122-9.

Additional file 1: Appendix A. West Virginia Department of Education Descriptive Writing Rubric. **Appendix B**. Smarter Balanced Narrative Writing Rubric. **Appendix C.** Samples of TA and PA on Descriptive Writing Tasks. **Appendix D**. Samples of TA and PA on Narrative Writing Tasks.

Availability of data and materials
Please contact the authors for data requests.

Declarations

Competing interests
The authors declare that they have no competing interests.

References
Abbott, H. P. (2002). *The Cambridge introduction to narrative.* Cambridge University Press.
Agbayahoun, J. P. (2016). Teacher written feedback on student writing: Teachers' and learners' perspectives. *Theory and Practice in Language Studies*, *6*(10), 1895–1904. https://doi.org/10.17507/tpls.0610.01.
Ames, H., Glenton, C. & Lewin, S. (2019). Purposive sampling in a qualitative evidence synthesis: a worked example from a synthesis on parental perceptions of vaccination communication. *BMC Med Res Methodol*, *19*(26). https://doi.org/10.1186/s12874-019-0665-4.
Ataie-Tabar, M., Zareian, G., Amirian, S. M. R., & Adel, S. M. R. (2019). A study of socio-cultural conception of writing assessment literacy: Iranian EFL teachers' and students' perspectives. *English Teaching & Learning*, *43*(4), 389–409. https://doi.org/10.1007/s42321-019-00035-0.
Azarnoosh, M. (2013). Peer assessment in an EFL context: Attitudes and friendship bias. *Language Testing in Asia*, *3*(1), 1–10.
Birjandi, P., & Hadidi Tamjid, N. (2012). The role of self, peer and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment & Evaluation in Higher Education*, *37*(5), 513–533. https://doi.org/10.1080/02602938.2010.549204.
Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*(2), 102–118. https://doi.org/10.1016/j.jslw.2007.11.004.
Brammer, C., & Rees, M. (2007). Peer review from the students' perspective: Invaluable or invalid? *Composition Studies*, *35*(2), 71–85.
Brown, D. (2004). *Language assessment: Principles and classroom practice.* Longman.
Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, *36*(4), 395–407.
Chala Bejarano, P. A., & Chapetón, C. M. (2013). The role of genre-based activities in the writing of argumentative essays in EFL. *Profile Issues in Teachers Professional Development*, *15*(2), 127–147.
Chang, C. Y., Lee, D. C., Tang, K. Y., & Hwang, G. J. (2021). Effect sizes and research directions of peer assessments: From an integrated perspective of meta-analysis and co-citation network. *Computers & Education*, *164*, 104123. https://doi.org/10.1016/j.compedu.2020.104123.
Chen, Y. S., & Su, S. W. (2012). A genre-based approach to teaching EFL summary writing. *ELT Journal*, *66*(2), 184–192. https://doi.org/10.1093/elt/ccr061.
Cheng, W., & Warren, M. (1999). Peer and teacher assessment of the oral and written tasks of a group project. *Assessment & Evaluation in Higher Education*, *24*(3), 301–314. https://doi.org/10.1080/0260293990240304.
Colognesi, S., & Deschepper, C. (2018). La relecture collaborative comme levier de soutien aux corrections des textes [Collaborative proofreading as a lever for rewriting and support for text corrections]. *Le Français Aujourd'hui*, *203*(4), 63–72. https://doi.org/10.3917/lfa.203.0063.
Deng, C., & Carless, D. R. (2010). Examination preparation or effective teaching: Conflicting priorities in the implementation of a pedagogic innovation. *Language Assessment Quarterly*, *7*(4), 285–302. https://doi.org/10.1080/15434303.2010.510899.
Derewianka, B. (2003). Trends and issues in genre-based approaches. *RELC Journal*, *34*(2), 133–154. https://doi.org/10.1177/003368820303400202.
Diab, N. M. (2015). Effectiveness of written corrective feedback: Does type of error and type of correction matter? *Assessing Writing*, *24*, 16–34. https://doi.org/10.1016/j.asw.2015.02.001.
Dolz, J., & Gagnon, R. (2008). Le genre du texte, un outil didactique pour développer le langage oral et écrit [the gender of the text, a didactic tool for developing oral and written language]. *Pratiques. Linguistique, Littérature, Didactique*, (137-138), 179–198. https://doi.org/10.4000/pratiques.1159.

Dressler, R., Chu, M. W., Crossman, K., & Hilman, B. (2019). Quantity and quality of uptake: Examining surface and meaning-level feedback provided by peers and an instructor in a graduate research course. *Assessing Writing*, *39*, 14–24. https://doi.org/10.1016/j.asw.2018.11.001.

Dumais, C., Lafontaine, L., & Pharand, J. (2017). J'enseigne et j'évalue l'oral: Pratiques effectives au 3e cycle du primaire ["I teach and evaluate oral communication": Actual practices at the 3rd cycle of primary school]. In J. F. de Pietro, C. Fisher, & R. Gagnon (Eds.), *L'oral aujourd'hui: Perspectives didactiques [Oral communication today: Didactic perspectives]*, (pp. 151–174). Presses Universitaires de Namur.

Farahani, A. A. K., Nemati, M., & Montazer, M. N. (2019). Assessing peer review pattern and the effect of face-to-face and mobile-mediated modes on students' academic writing development. *Language Testing in Asia*, *9*(1), 1–24.

Fathi, J., & Khodabakhsh, M. R. (2020). Self-assessment and peer-assessment in writing course of Iranian EFL students: An investigation of writing anxiety. *International Journal of English Language & Translation Studies*, *8*(1), 88–96.

Freedman, A., & Medway, P. (1994). *Learning and teaching genre*. Boynton/Cook.

Gagnon, R., De Pietro, F., & Fisher, C. (2017). Introduction. In J. F. de Pietro, C. Fisher, & R. Gagnon (Eds.), *L'oral aujourd'hui: Perspectives didactiques [Oral communication today: Didactic perspectives]*, (pp. 11–40). Presses Universitaires de Namur.

Hanjani, A. M., & Li, L. (2014). Exploring L2 writers' collaborative revision interactions and their writing performance. *System*, *44*, 101–114. https://doi.org/10.1016/j.system.2014.03.004.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487.

Horverak, M. O. (2016). Students' and teachers' perceptions on writing instruction inspired by genre-pedagogy and systemic functional linguistics. *Proceedings of CECIL's*, *5*, 58–73.

Hyatt, D. F. (2005). Yes, a very good point!: A critical genre analysis of a corpus of feedback commentaries on master of education assignments. *Teaching in Higher Education*, *10*(3), 339–353. https://doi.org/10.1080/13562510500122222.

Hyland, K. (2003). *Writing and teaching writing in second language writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667251.

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, *16*(3), 148–164. https://doi.org/10.1016/j.jslw.2007.07.005.

Jafarpur, A. (1991). Can naïve efl learners estimate their own proficiency? *Evaluation and Research in Education*, *5*(3), 145–157. https://doi.org/10.1080/09500799109533306.

Jaubert, M. (2007). *Langage et construction de savoirs à l'école, un exemple en sciences [language and knowledge building in schools, an example in science]*. Presses Universitaires de Bordeaux.

Jones, I., & Alcock, L. (2014). Peer assessment without assessment criteria. *Studies in Higher Education*, *39*(10), 1774–1787. https://doi.org/10.1080/03075079.2013.821974.

Joordens, S., Pare, D., & Pruesse, K. (2009). *Peerscholar: An evidence-based online peer assessment tool supporting critical thinking and clear communication*. Paper presented at the Proceedings of the 2009 International Conference on e-Learning.

Karimi, M. N., & Asadnia, F. (2015). EFL teachers' beliefs about oral corrective feedback and their feedback-providing practices across learners'proficiency levels. *Journal of Teaching Language Skills*, *34*(2), 39–68.

Labov, W. (1997). Some further steps in narrative analysis. *Journal of Narrative and Life History*, *7*, 395–415.

Lahuerta, A. C. (2018). Study of accuracy and grammatical complexity in EFL writing. *International Journal of English Studies*, *18*(1), 71–89. https://doi.org/10.6018/ijes/2018/1/258971.

Lam, R. (2010). A peer review training workshop: Coaching students to give and evaluate peer feedback. *TESL Canada Journal*, *27*(2), 114–114. https://doi.org/10.18806/tesl.v27i2.1052.

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

Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing*, *27*, 1–18. https://doi.org/10.1016/j.jslw.2014.10.004.

Li, L. (2017). The role of anonymity in peer assessment. *Assessment & Evaluation in Higher Education*, *42*(4), 645–656. https://doi.org/10.1080/02602938.2016.1174766.

Liao, H. C. (2016). Enhancing the grammatical accuracy of EFL writing by using an AWE-assisted process approach. *System*, *62*, 77–92. https://doi.org/10.1016/j.system.2016.02.007.

Liu, N. F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, *11*(3), 279–290. https://doi.org/10.1080/13562510600680582.

Liu, X., & Li, L. (2014). Assessment training effects on student assessment skills and task performance in a technology-facilitated peer assessment. *Assessment & Evaluation in Higher Education*, *39*(3), 275–292. https://doi.org/10.1080/02602938.2013.823540.

Liu, Y., & Huang, J. (2020). The quality assurance of a national English writing assessment: Policy implications for quality improvement. *Studies in Educational Evaluation*, *67*, 100941. https://doi.org/10.1016/j.stueduc.2020.100941.

Lou-Conlin, M. (1998). *Patterns: A short prose reader*.

Lu, J., & Law, N. (2011). Online peer assessment: Effects of cognitive and affective feedback. *Instructional Science*, *40*(2), 257–275.

Luquin, M., & Garcia Mayo, M. D. P. (2021). Exploring the use of models as a written corrective feedback technique among EFL children. *System*. https://doi.org/10.1016/j.system.2021.102465.

Man, D., Xu, Y., & O'Toole, J. M. (2018). Understanding autonomous peer feedback practices among postgraduate students: A case study in a Chinese university. *Assessment & Evaluation in Higher Education*, *43*(4), 527–536. https://doi.org/10.1080/02602938.2017.1376310.

Mao, S. S., & Crosthwaite, P. (2019). Investigating written corrective feedback:(Mis) alignment of teachers' beliefs and practice. *Journal of Second Language Writing*, *45*, 46–60. https://doi.org/10.1016/j.jslw.2019.05.004.

Matsuno, S. (2009). Self-, peer- , and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, *26*(1), 75–100. https://doi.org/10.1177/0265532208097337.

Min, H. T. (2016). Effect of teacher modeling and feedback on EFL students' peer review skills in peer review training. *Journal of Second Language Writing*, *31*, 43–57. https://doi.org/10.1016/j.jslw.2016.01.004.

Mirador, J. F. (2000). A move analysis of written feedback in higher education. *RELC Journal*, *31*(1), 45–60. https://doi.org/10.1177/003368820003100103.

Mohamadi, Z. (2018). Comparative effect of online summative and formative assessment on EFL student writing ability. *Studies in Educational Evaluation*, *59*, 29–40. https://doi.org/10.1016/j.stueduc.2018.02.003.

NBCT Office of Assessment West Virginia Department of Education (2015). *Artificial intelligence scoring of student essays: West Virginia's experience*.

Patri, M. (2002). The influence of peer feedback on self- and peer-assessment of oral skills. *Language Testing*, *19*(2), 109–131. https://doi.org/10.1191/0265532202lt224oa.

Plakans, L., Gebril, A., & Bilki, Z. (2019). Shaping a score: Complexity, accuracy, and fluency in integrated writing performances. *Language Testing*, *36*(2), 161–179. https://doi.org/10.1177/0265532216669537.

Pope, N. (2001). An examination of the use of peer rating for formative assessment in the context of the theory of consumption values. *Assessment & Evaluation in Higher Education*, *26*(3), 235–246. https://doi.org/10.1080/02602930120052396.

Pourdana, N., Nour, P., & Yousefi, F. (2021). Investigating metalinguistic written corrective feedback focused on EFL learners' discourse markers accuracy in mobile-mediated context. *Asian-Pacific Journal of Second and Foreign Language Education*, *6*(1), 1–18.

Rassaei, E. (2013). Corrective feedback, learners' perceptions, and second language development. *System*, *41*(2), 472–483. https://doi.org/10.1016/j.system.2013.05.002.

Rezvani, P., Aqdam, S. K., & Saeidi, M. (2014). *The effect of genre-based teaching upon EFL writing achievement*. Paper presented at the FLLT Conference Proceedings by Language Institute of Thammasat University.

Saito, H. (2008). EFL classroom peer assessment: Training effects on rating and commenting. *Language Testing*, *25*(4), 553–581. https://doi.org/10.1177/0265532208094276.

Saito, H., & Fujita, T. (2004). Characteristics and user acceptance of peer rating in EFL writing classrooms. *Language Teaching*, *8*(1), 31–54. https://doi.org/10.1191/1362168804lr133oa.

Sari, D. M. M. (2019). An overview of genre based approach in EFL writing class. *Journal of English Education (JournE)*, *1*(1), 31–40.

Saslow, J., & Ascher, A. (2006). *A. Top notch 3B. English for today's world*. Pearson Education.

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist*, *32*(4), 195–208. https://doi.org/10.1207/s15326985ep3204_1.

Sermsook, K., Liamnimitr, J., & Pochakorn, R. (2017). The impact of teacher corrective feedback on EFL student writers' grammatical improvement. *English Language Teaching*, *10*(10), 43–49. https://doi.org/10.5539/elt.v10n10p43.

Seror, J. (2011). Alternative sources of feedback and second language writing development in university content courses. *The Canadian Journal of Applied Linguistics*, *14*(1), 118–143.

Shen, B., Bai, B., & Xue, W. (2020). The effects of peer assessment on learner autonomy: An empirical study in a Chinese college English writing class. *Studies in Educational Evaluation*, *64*, 100821. https://doi.org/10.1016/j.stueduc.2019.100821.

Shih, R. C. (2011). Can web 2.0 technology assist college students in learning English writing? Integrating Facebook and peer assessment with blended learning. *Australasian Journal of Educational Technology*, *27*(5).

Smarter Balanced Assessment Consortium (2012). *Preformance task writing rubric*. Retrieved November 10, 2020 from https://commoncore.tcoe.org/content/public/doc/smarter_balanced_3-8_narr_writing_rubric.pdf

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

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.

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, *76*(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3.

Tian, L., & Zhou, Y. (2020). Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context. *System*, *91*, 102247. https://doi.org/10.1016/j.system.2020.102247.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes: Harvard University press*. https://doi.org/10.2307/j.ctvjf9vz4.

Wang, L., Lee, I., & Park, M. (2020). Chinese university EFL teachers' beliefs and practices of classroom writing assessment. *Studies in Educational Evaluation*, *66*, 100890. https://doi.org/10.1016/j.stueduc.2020.100890.

Wang, W. (2014). Students' perceptions of rubric-referenced peer feedback on EFL writing: A longitudinal inquiry. *Assessing Writing*, *19*, 80–96. https://doi.org/10.1016/j.asw.2013.11.008.

William, D. (2018). *Feedback: At the heart of - but definitely not all of - formative assessment*. Paper presented at the The Cambridge handbook of instructional feedback.

Xu, Y. T., & Liu, J. (2010). Research on L2 Writing Feedback Based on Anonymous Written Feedback. *Foreign Language Teaching*, *3*, 44–49.

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*, 179–200.

Yang, M., & Carless, D. (2013). The feedback triangle and the enhancement of dialogic feedback processes. *Teaching in Higher Education*, *18*(3), 285–297. https://doi.org/10.1080/13562517.2012.719154.

Yelland, C. (2011). A genre and move analysis of written feedback in higher education. *Language and Literature: International Journal of Stylistics*, *20*(3), 218–235. https://doi.org/10.1177/0963947011413563.

Yilmaz, Y. (2013). The relative effectiveness of mixed, explicit and implicit feedback in the acquisition of English articles. *System*, *41*(3), 691–705. https://doi.org/10.1016/j.system.2013.07.020.

Yu, S. (2020). Giving genre-based peer feedback in academic writing: Sources of knowledge and skills, difficulties and challenges. *Assessment & Evaluation in Higher Education*, *46*(1), 36–53. https://doi.org/10.1080/02602938.2020.1742872.

Yu, S., Jiang, L., & Zhou, N. (2020). Investigating what feedback practices contribute to students' writing motivation and engagement in Chinese EFL context: A large scale study. *Assessing Writing*, *44*, 100451. https://doi.org/10.1016/j.asw.2020.100451.

Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Language Teaching*, *49*(4), 461–493. https://doi.org/10.1017/s0261444816000161.

Zareei, A. (2009). The effects of teaching genre moves on EFL learners performance in letter writing. *Pazhuhesh-e Zabanha-ye Khareji*, *49*, 43–64.

Zarifi, A. (2017). Iranian EFL learners' reaction to teacher's written corrective feedback. *International Journal of Applied Linguistics and English Literature*, 6(3), 254–261. https://doi.org/10.7575/aiac.ijalel.v.6n.3p.254.

Zenouzagh, Z. M. (2020). Syntactic complexity in individual, collaborative and E-collaborative EFL writing: Mediating role of writing modality, L1 and sustained development in focus. *Educational Technology Research and Development*, 68(6), 2939–2970. https://doi.org/10.1007/s11423-020-09818-w.

Zhao, C. G., & Liao, L. (2021). Metacognitive strategy use in L2 writing assessment. *System*, 102472.

Zhao, H. (2014). Investigating teacher-supported peer assessment for EFL writing. *ELT Journal*, 68(2), 155–168. https://doi.org/10.1093/elt/cct068.

Zhao, H. (2018). Exploring tertiary English as a foreign language writing tutors' perceptions of the appropriateness of peer assessment for writing. *Assessment & Evaluation in Higher Education*, 43(7), 1133–1145. https://doi.org/10.1080/02602938.2018.1434610.

Zhao, H., & Zhao, B. (2020). Co-constructing the assessment criteria for EFL writing by instructors and students: A participative approach to constructivelyaligning the CEFR, curricula, teaching and learning. *Language Teaching Research*. https://doi.org/10.1177/1362168820948458.

Publisher's Note