

A PROPOSAL OF FORMATIVE ASSESSMENT IN EFL TEACHING AND LEARNING: ONLINE WRITING AND PEER-REVIEW ACTIVITIES

Kátia Muck^{1*}

¹Centro Universitário Facvest – UNIFACVEST, Lages, SC, Brasil

Bill Cope^{2**}

²University of Illinois at Urbana-Champaign, United States of America

Abstract

Technological advances have set new roles and modified the relations among those who engage in the teaching-learning process, demanding new and efficient types of formative assessment. Guided by the research question “What kind of peer-feedback was used for the development of the students’ written case studies?”, this study examines 72 reviews to 29 papers developed by 31 learners taking a master’s course, employing the CGScholar platform. Statistical analyses on reviewers’ reliability (ICC agreement and consistency) and qualitative analysis on the descriptive feedback were conducted. The results from the quantitative analyses show high reliability among reviewers; and the qualitative analysis suggests that these activities can be employed in EFL learning to enhance learning opportunities, employ higher cognitive processes, and to teach large groups of students, either remote or face-to-face, without increasing management time.

Keywords: EFL; Online Writing, Online Peer-Review, Formative Assessment; CGScholar

* Professor, Postgraduate Program on Transcultural Practices, Centro Universitário Facvest-UNIFACVEST. Leader of a research group on the integration of new technologies and the teaching-learning processes. CV: <http://lattes.cnpq.br/6203189529242782> E-mail: katiamuck@gmail.com. ORCID: <https://orcid.org/0000-0001-9702-9026>.

** Professor, Department of Education Policy, Organization & Leadership, College of Education, University of Illinois at Urbana-Champaign. Website: <http://wwcope.com>. E-mail: billcope@illinois.edu. ORCID: <https://orcid.org/0000-0003-2129-395X>.



1. Introduction

The affordances of the technological development have changed social practices. In this context, learning also presents new genres and practices (Burbules, 2010). According to Haythornthwaite, (2010, p. 37), “[n]ew technologies forge new relations and new roles for participants”. This author also adds that “this is highly evident in the way online spaces are transforming educational and authoritative practice” (Haythornthwaite, 2010, p. 37), such as the way knowledge production, distribution and consumption have shifted from teacher-focused to ubiquitous learning. Students have become protagonists of their learning process, and this requires efficient types of formative assessment.

The use of peer-activities in education has increased over the years, especially due to technological advancements. Dambros (2020), for example, employed WhatsApp to develop collaborative writing activities in Portuguese with her junior high school students. Likewise, the teaching and learning of English as a foreign Language (EFL) has explored a range of technological tools to foster peer-activities to enhance learning. For instance, researchers have employed Google Docs/Drive (Slavkov, 2015; Jeong, 2016; Seyyedrezaie et. al., 2016; Alharbi, 2019) and Edmodo (Paker & Dogan, 2021) to develop collaborative writing and editing in EFL, Facebook to provide teacher written feedback (Elfiza, Reszki & Nopita, 2021) and to promote English language learning and active online participation through peer-writing discussions (Al Qunayeer, 2020), and Wiki for peer-feedback in EFL essay writing (Abri-Al, Baimani-Al & Bahlani, 2021).

However, except for Edmodo, all the above-mentioned tools were not conceived specifically for educational purposes. Therefore, the learning objectives may become shallow due to the limitations of the technological resource. Moreover, activities such as peer-feedback may become a burden for teachers, as they require an incredible amount of managerial time (Kern et al., 2002; Kern et al., 2003).

To address this gap, this article suggests the use of the CGScholar multimodal platform to develop writing skills in EFL, peer-feedback, self-assessment, teacher-feedback, social interaction, grading, and the distribution of the written production. Due to space constraints, this report will focus on writing activities, peer-feedback, and self-assessment by reporting research findings on the appropriateness of this resource. Although the investigated participants were not language learners, we advocate the affordances of CGScholar both for learning *per se* in any subject and for teachers, as it is resourceful, effective to manage, and it allows the development of several learning objectives. The present study was guided by the following research question: What kind of peer-feedback was used for the development of the students’ written case studies?

Thus, after this introduction, this paper provides a brief review of literature on peer-review and on the theory behind CGScholar. Subsequently, it presents the method, describing the context of investigation, the participants, the materials and the peer-feedback activity, as well as the procedures for data collection and analyses. Then, it delivers the results and discussion followed by the final remarks.

2. Review of Literature

This section offers an overview on peer-review and on the New Learning theory for the following reasons: 1) the investigated course follows the principles of this theory; 2) the online environment CGScholar was developed according to the fundamentals of this theory; and 3) this research adopts this theory as it may provide social change through educational environments such as the peer-review activity developed in the present investigation.

It is relevant to become aware of new educational demands regarding literacy practices for all those involved in the teaching-learning process and in multiliteracies (Heberle & Abreu, 2011). Butin (2012) advocates for the need of a model of formative feedback instead of summative feedback in e-learning. Cope and Kalantzis (2013) illuminate this e-learning scenario both in the theoretical and in the practical dimensions. In the former, they indicate *New Learning* as a new direction for education. In this paradigm, education is seen as a constant co-construction of knowledge that takes place everywhere at any time. Moreover, it places the teacher as a facilitator and the students as autonomous agents responsible for their own construction/consumption of knowledge. This new relation with knowledge demands new ways of developing the teaching-learning social practice in online and blended learning environments.

In order to address this need, Cope and Kalantzis (2013) coordinate a project assembling a multidisciplinary team of professionals (educational researchers, software engineers, computer scientists, computational linguists, and psychometricians) to develop a learning platform named CGScholar. As stated by the authors (2013, p. 333), “the *Scholar* intervention is an attempt to reframe the relations of knowledge and learning, recalibrating traditional modes of pedagogy in order to create learning ecologies which are more appropriately attuned to our times”.

In this line of reasoning, Kalantzis and Cope (2012) present an agenda for new learning and assessment (Figure 1), which proposes seven openings for educational transformation: ubiquitous learning, active knowledge production, multimodal knowledge representation, recursive feedback, collaborative intelligence, and differentiated learning. Although these openings are already known in the educational theories or practices, the authors’ research on the subject “has attempted to explore ways in which what [they] have termed ‘social knowledge’ technologies might make each of these ideas easier to realize” (Cope & Kalantzis, 2013, p. 354).

Figure 1: Seven practical openings for educational transformation

Source: Retrieved from Cope and Kalantzis (2013, p. 333, *Figure 1. Seven openings, seven affordances*)

With education transformation, the focus is on the process rather than on the product. “Assessment is at the heart of formal higher education” (Gikandi et al., 2011, p. 2234). Therefore, more consideration should be given to formative assessment rather than summative assessment. The former supports learning and the latter provides validation and accreditation (Kollar & Fischer, 2010).

Gikandi et al. (2011) define online formative assessment as “the application of formative assessment within learning online and blended settings where the teacher and learners are separated by time and/or space and where a substantial proportion of learning/teaching activities are conducted through web-based ICT” (p. 2337). These authors conducted a systematic qualitative review of literature to understand: 1) “how formative assessment support learners in developing domain content knowledge and professional skills in an online environment”, and “core assessment concepts of validity and reliability as they occur in online contexts” (Gikandi et al., 2011, p. 2334). They reviewed 91 articles published until 2010 by employing the following search terms: “online assessment, online formative assessment, innovative assessment, assessing online learning, assessment in higher education, online formative assessment in higher education and alternative assessment” (Gikandi et al., 2011, p. 2334). Eighteen key studies were selected to be reviewed.

According to the authors, fundamental issues of assessment in online contexts are validity, reliability, and dishonesty. They define validity within the context of online formative assessment as “the degree to which the assessment activities and processes promote further learning” (Gikandi et al., 2011, p. 2338). They identified that characteristics such as authenticity of assessment activities, effective formative feedback, multidimensional perspectives, and learner support are associated to the mentioned validity. In turn, the authors define reliability within the context of online formative assessment as the “degree to which what is assessed is dependable or sufficient to measure the level of knowledge structure being developed (the desired learning outcomes)” (Gikandi et al., 2011, p. 2339). The authors identified that the following characteristics relate to

reliability: opportunities for documenting and monitoring evidence of learning, (2) multiple sources of evidence of learning and (3) explicit clarity of learning goals and shared meaning of rubrics. Finally, dishonesty, in this context, “relates to students truly owning their work, depends on the degree of inherent validity and reliability. This implies that dishonesty can be minimized through enhancing the identified aspects of validity and reliability” (Gikandi et al., 2011, p. 2341).

The online activities investigated in this study address the abovementioned issues of validity and reliability. Gikandi et al. (2011) state that by addressing these issues of validity, reliability and dishonesty “online formative assessment can function as an innovative pedagogical strategy through facilitating the following opportunities: (1) formative and immediate feedback, (2) engagement with critical learning processes, and (3) promoting equitable education” (p. 2344) by attending students’ individual differences.

Regarding formative assessment by peers, the authors conclude that “online formative assessment can provide learners with authentic, collaborative, and reflective learning environments to share learning experiences and dissonance of practice. These experiences emulate real professional communities of practice; thus promoting learner ability to apply knowledge to their own practice” (Gikandi et al., 2011, p. 2344). Moreover, “online settings can offer enhanced opportunities to provide more detailed and clearly written feedback that is integrated within student work” (Wolsey, 2008, in Gikandi et al., 2011, p. 2345). Besides, according to Nicol and Macfarlane (2006, cited in Gikandi et al., 2011, p. 2346), effective formative feedback: 1) helps clarify what good performance is (goals, criteria, expected standards); 2) facilitates the development of self-assessment (reflection) in learning; 3) delivers high quality information to students about their learning; 4) encourages teacher and peer dialog around learning; 5) encourages positive motivational beliefs and self-esteem; 6) provides opportunities to close the gap between current and desired performance; and 7) provides information to teachers that can be used to help shape teaching.

Kollar and Fischer (2010) defend the idea that peer assessment is still in its “adolescent” stage, and, as inherent to this stage, it is in search of its identity and its place in research fields. The authors describe the typical structure of peer assessment as the following: task performance, feedback provision, feedback reception, and revision. They state that simple engagement in this process does not guarantee that learning takes place. According to them, “when learning is seen as high-level change in an individual’s knowledge base, then, to make peer assessment a successful enterprise, it is necessary that high-level cognitive processing occurs” (Kollar & Fischer, 2010, p. 6).

They provide examples of actions that might facilitate high-level cognitive and discursive processing during each step of peer feedback activities. For instance, planning, reviewing, explaining, arguing, and questioning are examples of high-level cognitive processes developed during task performance, which is the step of the writing of the case study in this investigation. Understanding, planning, and monitoring are examples of high-level cognitive processes developed during

feedback provision. “For B’s feedback to facilitate A’s learning, B not only needs to process A’s first product, but also show planning and monitoring concerning how to formulate feedback in a way that A can benefit from it” (Kollar & Fischer, 2010, p. 6). Regarding the high-level processes students engage in during feedback reception, according to the authors, A will examine the received feedback, compare the comments with the original performed task (case study), and decide whether to employ the suggestions on the writing of the next version. This process is successful, according to the authors, when the feedback presents good quality by providing good arguments. Finally, comparison processes are fostered during revision, as students compare the first version, the feedback and the prospective revised version.

Similarly, Yu and Wu (2013) explain the cognitive processes that are mobilized during peer-assessment activities.

Assessing the relative quality and merits of the examined work encourages students to engage in critical thinking. In addition, both social and argumentation skills as well as substantial knowledge in the applied area are required to enable comments to be accepted by peers. Also, when observing peers’ work, students are likely to be alerted to problems that may exist in their own work and be prompted to make necessary modifications. On the other hand, when students receive feedback from assessors, the comments provided may cause cognitive conflict and direct students to deal with their existing cognitive defects. Knowledge structuring and re-structuring are cultivated through various cognitive and discursive processes (such as deeper elaboration of materials, self-reflection, comparison, clarification, adjustment, and so on). (Yu & Wu, 2013, p. 333)

The authors also observe that as students tend “to be within or near each other’s zone of proximal developments, peers’ comments may be more easily understood by learners than instructors” (Ammer, 1998; Fallows & Chandramoham, 2001; cited in Yu & Wu, 2013, p. 333).

Noroozi, Biemans, and Mulder (2016) came to analogous conclusions after analyzing the results of a research they conducted with 189 undergraduate BSc students in the Netherlands. They investigated the relations between peer feedback learning processes and outcomes during a peer-feedback activity that aimed at improving students’ performance on writing essays. Results demonstrated that students who provided high-quality feedback performed better on their final essay than students who provided poor feedback. The same relation happened with students that received high-quality feedback versus students that received poor quality feedback. According to the authors, this is because constructing and supporting arguments along with considering multiple perspectives demand complex cognitive processes. The same complex processes occur when students analyze and evaluate writings from their peers (Noroozi et al., 2016, p. 29).

A similar relation was identified by Pol, Berg, Admiraal, and Simons (2008), although investigating students’ views. They investigated “the relationships between the nature of feedback, its reception by the receiver, and its consecutive use in the revision of students’ texts” (Pol et al., 2008, p. 1805). Data were

collected during six months on peer feedback activities on several assignments from a group of 27 college students in the Netherlands. No significant results were found on the relationship between the nature of feedback and revision of products. However, results on the relationship between the reception of feedback and the use of feedback demonstrated that the more valuable students considered the feedback the more they employed it on the revision of their writing product.

On another study on writing, Cheng, Liang, and Tsai (2015) conducted a research on online peer assessment with 47 undergraduate students of Biology, in Taiwan, to investigate the role of feedback on students' writings. Their objective was to understand what and how peer-review may influence learning. Students went through three rounds of peer-review, reviewing five reports in each round. The students and the teacher had to provide descriptive feedback in five dimensions (knowledge, suitability, correctness, creativity, and overall) as well as a score from 1 to 7 for each dimension. In all rounds, the correlation coefficient r between the peer and teacher scores was significant, except for two dimensions in round one. The 705 messages of descriptive feedback were categorized into: Affective (Supporting; Opposing); Cognitive (Direct correction; Personal opinion; Guidance); Metacognitive (Evaluating; Reflecting); and Irrelevant comments. Results show that while the number of Affective feedback increased across the three rounds, the number of Cognitive and Metacognitive feedback, in general, decreased. However, "cognitive feedback messages were more helpful for these students' writing learning gains as compared with affective feedback (either positive or negative comments) and metacognitive feedback" (Cheng et al., 2015, p. 82).

Yang (2016) investigated 24 graduate English as foreign language (EFL) students of a master's program of EFL teaching and business communication in Taiwan. Their objective was to scrutinize academic knowledge transformation and construction during peer feedback activity on writing summary by using a computer-supported collaborative learning (CSCL) system. Students were separated into two groups: one experimental group and one control group. The former provided online peer feedback and the latter provided paper-based peer feedback. Students' perceptions on the matter were also investigated through surveys with open-ended questions.

The results show that students from the experimental group outperformed students from the control regarding the final text density. Moreover, the results suggest that "transforming and constructing academic knowledge through online summary writing and peer feedback helps graduate students raise their language awareness and critical thinking. By providing and receiving useful summary revisions from peers, the graduate students were able to recognize the key elements in well-organized academic texts and clarify illogical sentences and text misunderstanding" (Yang, 2016, p. 697). Concerning students' perceptions towards academic knowledge transformation and construction with the peer-review activity, most students responded that they enjoy providing feedback to peers because they can learn from each other online (12 out of 13 respondents), and that by giving feedback they are able to view other peers' summaries and compare them with their own (10 out of 13 respondents) (Yang, 2016, p. 696).

As noted by Kollar and Fischer (2010), authors employ different terminology to describe the same activity, such as “peer assessment”, “peer revision” and “peer feedback”. Besides these terms, this article also uses the term “peer-review”. All of them are employed interchangeably in this article.

3. Method

This study uses mixed methods design that involves quantitative and qualitative analyses to unveil the impact of peer review processes. This research endeavor was approved by the Institutional Review Board of the University of Illinois at Urbana-Champaign under IRB# 14.439 and is guided by the following research question: What kind of peer feedback was used for the development of the students’ written case studies?

3.1 Context of Investigation

The participants were taking “EPSY 408 – Learning and Human Development with Technologies”, a course for the Master’s degree program in Education offered completely online by the University of Illinois Urbana-Champaign. Its objective was to develop an understanding of theories of learning and relate them to educational technology. It was taught in eight weeks, in 2014, with the following schedule: Week 1: Introduction; Week 2: Behaviorism and Conditioned Response; Week 3: Notions of Innate Intelligence; Week 4: Constructivism; Week 5: Neuroscience; Week 6: The Social Mind; Week 7: Distributed Cognition; Week 8: Communities of Practice. The course workload was the following: 1) Writing of Work 1 and Work 2; 2) Peer reviewing three other participants’ works (in each work: 1 and 2). Revising their work considering the peer review comments and writing a self-review; 4) Commenting on the weekly discussion topic updates; 4) Posting at least seven weekly updates, reading others’ updates, and commenting on three of them; 5) Participating in the weekly 1.5-hour online synchronous encounters every Monday. Activities 1 to 4 were developed in Scholar and activity 5 in Adobe Connect.

3.2 The Participants

A total of 31 learners participated in this investigation. Nineteen are female and 12 are male, being 26 from 23 to 49 years old and 2 above 50 years old. Regarding their formal education background, one holds a doctorate’s degree, eight hold a master’s degree, and 19 hold bachelor’s degree. Also, 26 reported having teaching experience: six participants have between 5-10 years and six between 15-23 years of teaching experience. Twenty-four participants are native speakers of English and two of Chinese. Some participants did not provide some demographic information.

3.3 Materials and Activity

The materials that provided data for this study are the pre- and post-course surveys, the participants’ written case studies (Works) and reviews. The pre-course survey investigated participants’ demographics, and post-course survey explored participants’ experiences with the peer review activity. Both multiple-choice Likert-5 scale and open-ended questions were employed.

The activities under analysis were regular tasks of the course. Learners had to: write a case study contemplating the six sections and following the established rubrics (Table 1); review case studies from three other peers; review their own case study following the rubrics; and revise it based on the feedback they received from their peers. Besides providing descriptive feedback for each criterion, reviewers had to numerically rate each section from zero to four, with these numbers having specific values for criteria.

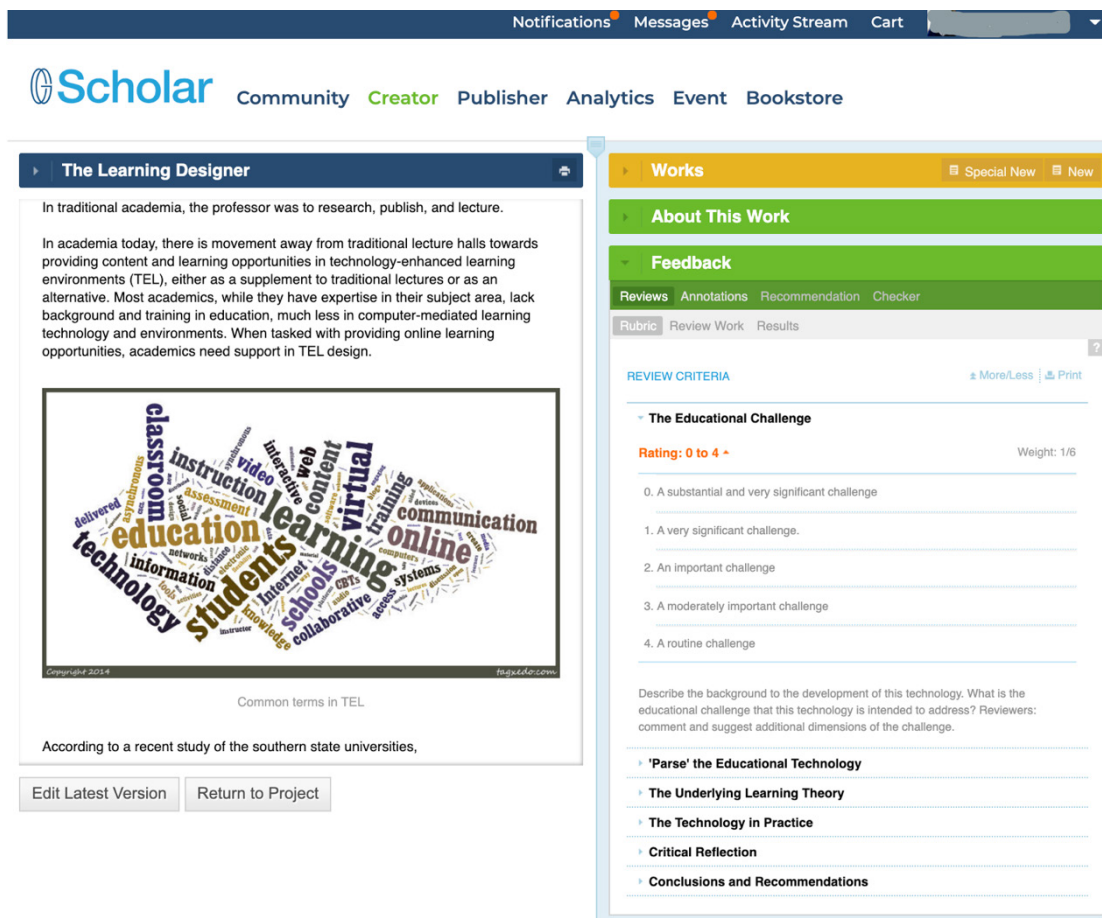
Table 1: Rubrics to write and review the Work

Sections of the Work	RUBRICS		
	For the Writer	For the Reviewer (qualitative feedback)	For the Reviewer (numerical rate)
The Educational Challenge	Describe the background to the development of this e-learning practice or technology. What is the educational challenge that the practice or technology is intended to address?	Comment and suggest additional dimensions of the challenge	4: A substantial and very significant challenge 3: A very significant challenge. 2: An important challenge 1: A moderately important challenge 0: A routine challenge
Parse' the Ecology	How does the e-learning practice or technology work? Explain its structure and function from the point of view of teachers and learners. How does the way the practice or technology works reflect its understanding of learning?	What else would you like to know?	4: Excellent information coverage 3: Good information coverage 2: Enough information 1: Just enough information 0: Not enough information
The Underlying Learning Theory	What is the underlying theory of learning that this educational practice technology reflects? Use the seven affordances framework to analyze this e-learning environment or technology	Comment and suggest possible additional theoretical perspectives.	4: Excellent connections made with learning theories 3: Good connections made with learning theories 2: Adequate connections made with learning theories 1: Very little or no connections made with learning theories 0: Insufficient connections made with learning theories
The Technology or Learning Process in Practice	Is the operation of the technology in practice and associated learning practices adequately described in the case study? Are concrete examples provided? Do they illustrate the way the environment's underlying theory of learning translates into practice?	If you believe more information is needed for a full interpretation and explanation of this e-learning environment, suggest what.	4: Excellent description 3: Very good description 2: Good description 1: Adequate description 0: Poor description
Critical Reflection	What are the relative strengths and weaknesses of the technology or practice from the perspectives of teaching and learning?	Provide comments and suggestions about other possible strengths and weaknesses that the author may not have considered.	4: Comprehensive analysis of strengths and weaknesses 3: Good analysis of strengths and weaknesses 2: Adequate analysis of strengths and weaknesses 1: Partial analysis of strengths and weaknesses 0: Minimal analysis of strengths and weaknesses
Conclusions and Recommendations	Do the conclusions and recommendations follow from the information and reasoning provided in the case study?	Comment and suggest conclusions and recommendations you might want to see added	4: Very clear conclusions and strong, comprehensive recommendations 3: Strong conclusions and recommendations 2: Partial conclusions and incomplete recommendations 1: Poor conclusions and limited recommendations 0: Unclear conclusions and impractical recommendations

Source: Elaborated with information from the Course Syllabus

These activities were performed in CGScholar, specifically in the space called *Creator*. Figure 2 shows a general view of this space where the Work named “The Learning Designer” was being developed. This multimodal space displays the space for writing the Work (left side) parallel to all the rubrics (right side) necessary to write and to review the Work (rubrics for this activity are in Table 1). Each section of the review criteria can be extended revealing the description of the criteria and the rating categories.

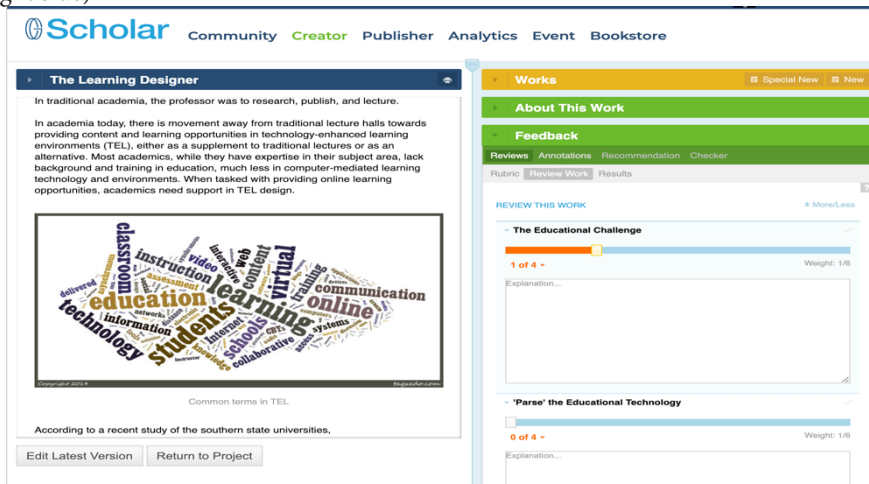
Figure 2: General view of the *Creator* space



Source: Partial screenshot of CGScholar, Creator, EPSY408 course

Muck (2015) and Muck and Sadki (2015) explore more features from CGScholar; due to space constraints, the focus here is on the features that foster feedback. Figure 3 exhibits the space to type in the qualitative feedback and to provide quantitative review by sliding the bar with the numbers to the right.

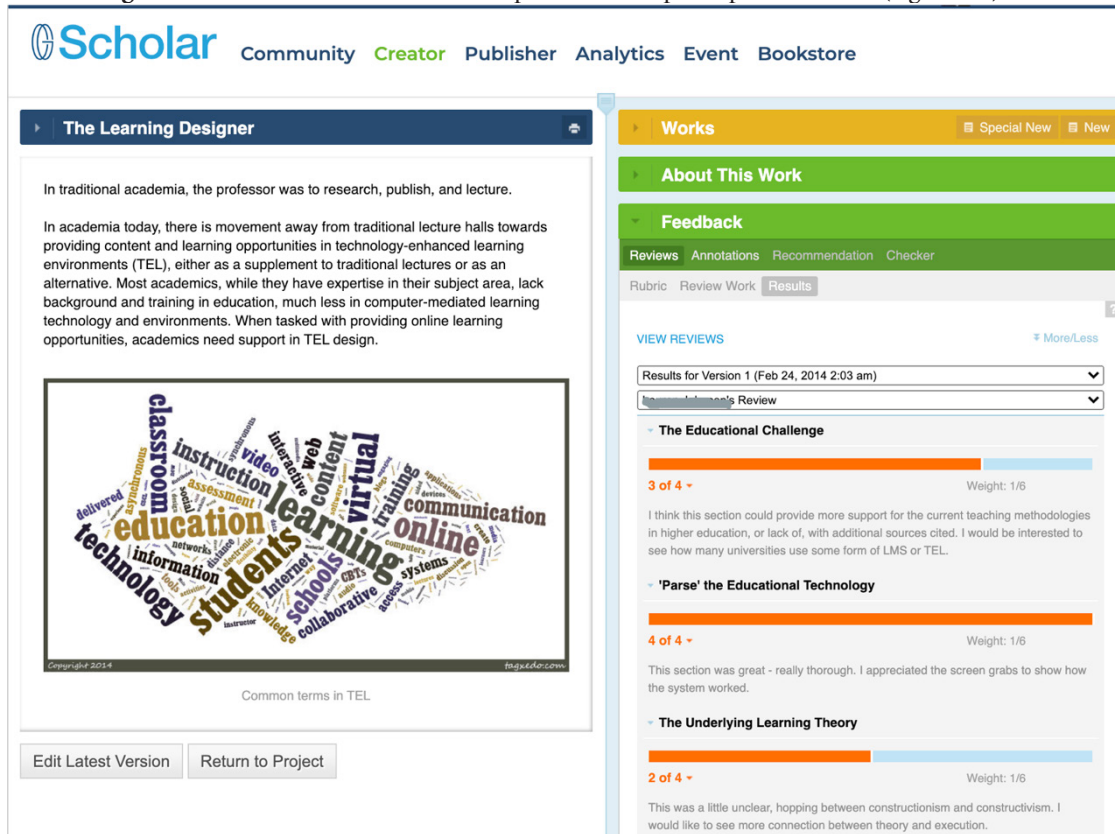
Figure 3: General view of *Creator*: space to provide qualitative and quantitative feedback (right side)



Source: Partial screenshot of CGScholar, Creator, EPSY408 course

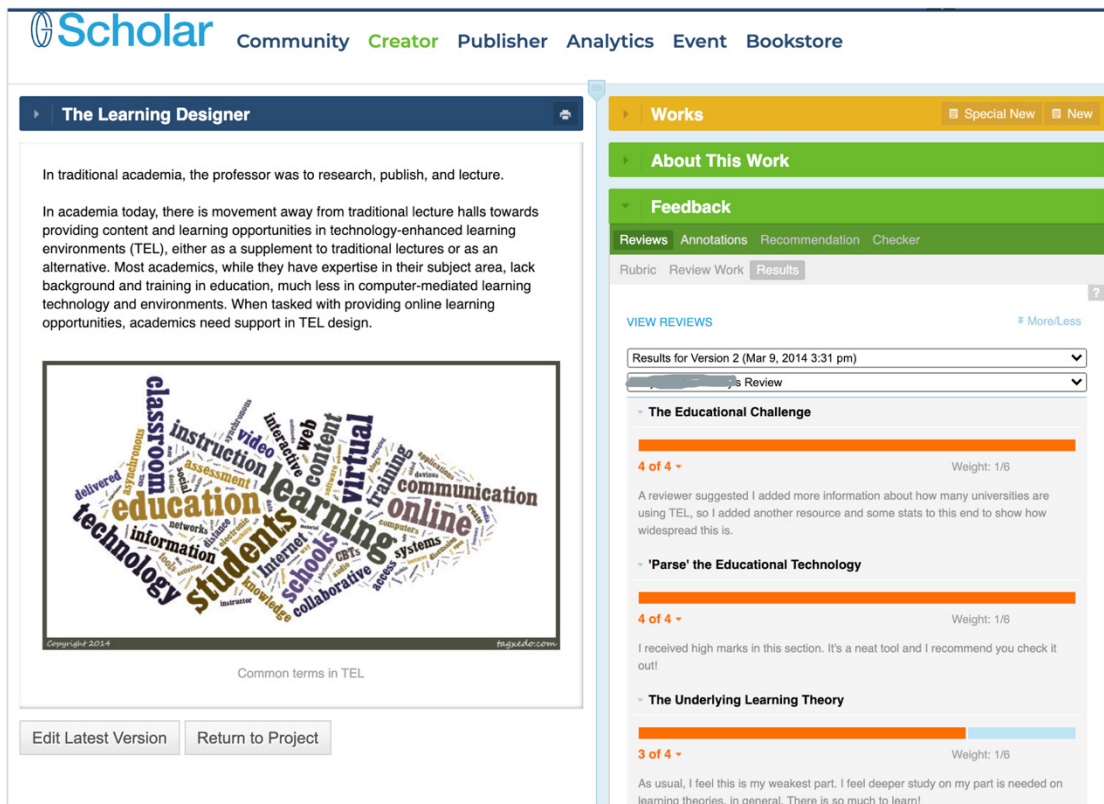
Figures 4 and 5 provide instances of feedback. The former shows the result of peer-review, and the latter displays an instance of a self-review after receiving the feedback from three peers.

Figure 4: General view of the Creator space: an example of peer feedback (right side)



Source: Partial screenshot of CGScholar, Creator, EPSY408 course

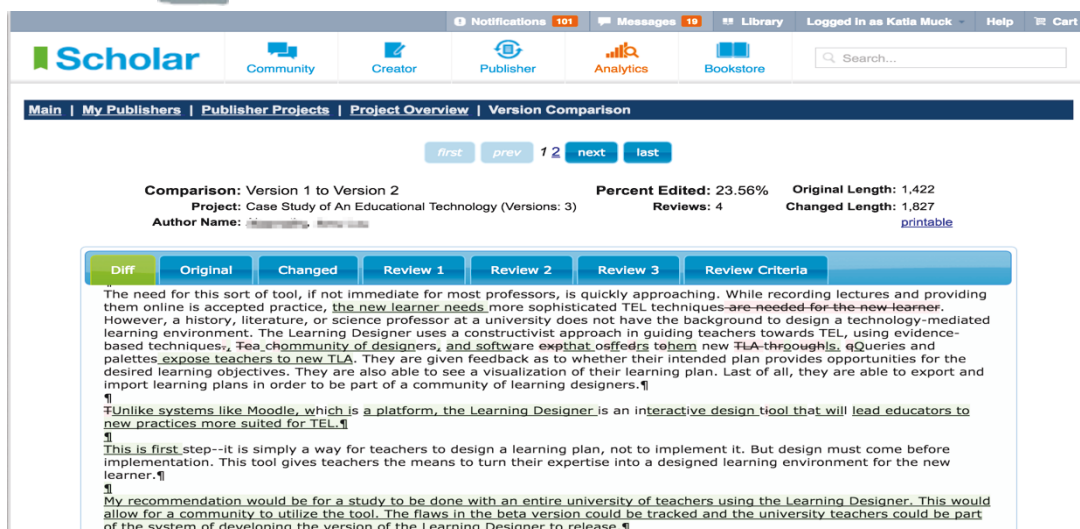
Figure 5: General view of the Creator space: an example of self-feedback (right side)



Source: Partial screenshot of CGScholar, Creator, EPSY408 course

CGScholar's *Analytics* is for facilitators to manage the peer-review activity. It allows the facilitator to track, for instance: the different versions that the participant writes, the submitted version to be reviewed, the reviewing criteria, the reviewer's feedback. Moreover, the facilitator can access a marked-up version indicating all the difference between the versions (Figure 6). This Figure shows an excerpt of what the learner edited. In total, s/he edited 23.56% of her case study by including information, as indicated by the green color in the excerpt, or by excluding information, as illustrated by the light pink color with the strikethrough effect.

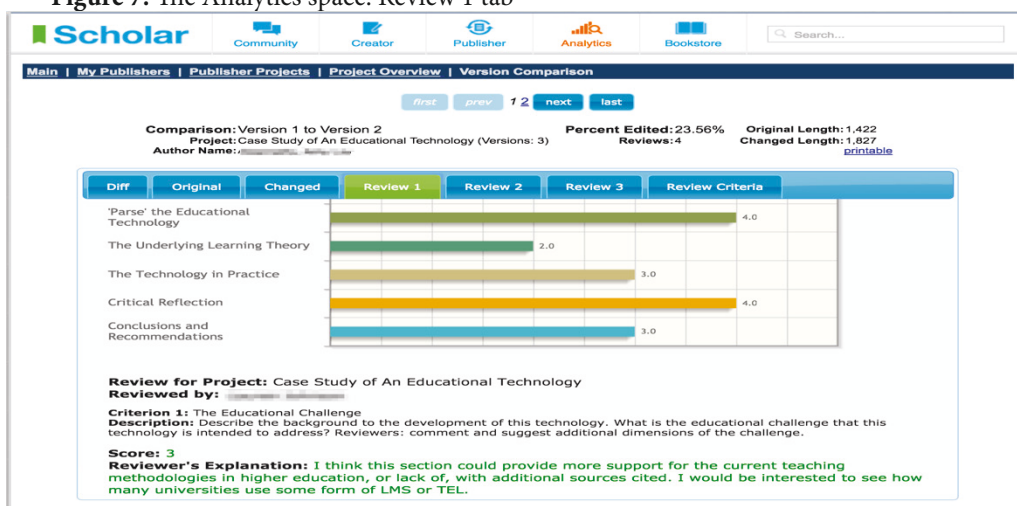
Figure 6: The Analytics space: Diff tab



Source: Partial screenshot of CGScholar, Analytics: EPSY408 course, manipulated in Skitch for redactions

Besides this marked up version, revealed by the first tab (“Diff”, in green) in Figure 6, the adjacent tab “Original” shows the writer’s original work, the tab “Changed” shows the revised work (without mark ups), the tab “Reviewer 1” shows the numeric feedback and the qualitative feedback that Reviewer 1 provided (Figure 7) and the same as the tabs of the other reviewers. The “Review Criteria” tab displays the review criteria. Furthermore, the facilitator can have an overview of the students’ achievements such as average number of words that each student wrote in the writing assignment, the percentage of editing each student did on their works, number of reviews s/he received and average grade, just to cite some features.

Figure 7: The Analytics space: Review 1 tab



Source: Partial screenshot of CGScholar, Analytics, EPSY408 course, manipulated in Skitch for redactions

Figure 7 exposes a part of the bar chart with the summary of the numeric feedback that Reviewer 1 gave to this work. In the sequence, it shows an excerpt of the descriptive feedback. It displays the criterion 1, the reviewer's score and the reviewer's explanation. The same sequence is available for the rest of the criteria.

3.4 Procedures for Data Collection and Analyses

Participants were recruited by e-mail with a link to take the pre-course survey. The post-course survey started in the last week of the course and the survey continued active for two weeks after the course ended. Survey was designed, distributed and organized employing Qualtrics. Data from the Works and the reviews were collected directly from Scholar.

All the data were exported to SPSS v.24. In the sequence, the survey was linked to the classroom data (the reviews). Each participant was assigned with an ID and their work products (drafts of works and peer review comments) and survey had their identities (names, e-mails, or other identifiable marks) removed.

This research employed quantitative and qualitative analyses. The former was employed to measure Reliability among reviewers. "The main defining characteristic of rater reliability is that scores by two or more raters [...] are consistent" (Dornyei, 2007, p. 57). Reliability, according to Silverman (p. 224, in: Dornyei, 2007, p. 57), "refers to the degree of consistency with which instances are assigned to the same category by different observers or by the same observer on different occasions". The more adequate tests to measure Reliability between reviewers, according to Deniszczwicz and Kern (2013) are the Intraclass Correlation Coefficient ICC (agreement) and the ICC (consistency), being the former more suitable than the latter. As they reveal Reliability from different perspectives, statistics for both types were run employing the statistics program SPSS v.24. It was employed the ICCs Alpha for Consistency and Absolute Agreement Types, with confidence interval of 95%.

With regard to the qualitative analysis, data was organized in categories that emerged from the data. The descriptive feedback provided by the reviewers was organized into each review criterion and, further, into the scores for each review criterion, in the following categories: 1) Additional Comments/Suggestions (AC/S): feedback with additional comments and suggestions related to the established criterion; 2) Additional Comments/Suggestions (AC/S-N): feedback with additional comments and suggestions that could improve the work, although not related to the established criterion; 3) Just Comments (JC): feedback that does not have potential to improve the writing; and 4) Unclear (U): feedback that was impossible to understand.

4. Results and Discussion

As detailed in the Method section, students engaged in an activity of blind peer reviewing each other's works. This Work is comprised by six sections and each one of them should receive a numeric feedback, which is a grade from zero

to four, as well as a descriptive feedback. Both types of feedback had to be based on the provided rubrics. Bearing these two types of feedback in mind, this section is organized as follows: Firstly, it presents the data from the numeric feedback and a statistic analysis regarding the reliability of the reviewers. In the sequence, it provides a qualitative analysis of the descriptive feedback from reviewers.

4.1. Quantitative approach

Regarding the reliability of the reviewers, Table 2 shows the results of the statistical analyses on reliability of the 72 reviews to the 29 papers developed in the course; each paper received two or three valid reviews, and two other papers received only one review each and were excluded from these analyses. For both ICC (*agreement*) and ICC (*consistency*), four cases present negative ICCs (S13, S7, S29, and S12) and all the others present positive ICCs, having two cases with ICCs equal to zero (S21 and S27).

Table 2: Intra Class Correlation results - ICC (*agreement*) and ICC (*consistency*) - for reviews in Scholar

ID	S13	S7	S29	S12	S21	S27	S16	S32	S8	S5	S30	S4	S9
Valid Reviews	2	2	3	3	2	2	3	2	3	3	3	3	3
ICC (agreement)	-2.723	-.667	-.176	-.122	.000	.000	.014	.048	.094	.096	.156	.286	.296
ICC (consistency)	-2.246	-1.500	-.207	-.105	.000	.000	.125	.143	.129	.277	.174	.441	.372

ID	S22	S26	S1	S28	S3	S25	S19	S6	S31	S10	S20	S18	S17	S23	S2	S15
Valid Reviews	3	2	3	2	3	2	3	2	2	3	2	3	2	2	2	2
ICC (agreement)	.310	.348	.435	.444	.545	.561	.585	.615	.646	.706	.713	.730	.783	.783	.800	.928
ICC (consistency)	.375	.471	.415	.429	.571	.604	.860	.769	.790	.800	.689	.758	.844	.750	.833	.928

Source: Table built with the individual results of the statistical analyses on reliability obtained with SPSS.

The ICC (*agreement*), which estimates the level of agreement among reviewers, varies from -2.723 to .928. Except for the first case (S13), the subsequent 12 cases (in the upper part of Table 2) exhibit ICC (*agreement*) (-.667 to .296) very similar to Locke, Silverman, and Spirduso (1998, p. 142)’s study (-.500 to .202, with three negative results and nine positive) with 12 graduate students engaging a similar peer review activity. Significant is the fact that the last 12 papers from the lower part of Table 12 (S3, S25, S19, S6, S31, S10, S20, S18, S17, S23, S2, and S15) present values for ICC (*agreement*) between .545 and .928, which is very near to the ideal level of agreement (that is 1). In other words, they present more

than 54% of agreement for each paper. It means, for example, that the reviewers of the paper S15 agree on 92.8% of the scores they gave for this paper.

In an analogous manner, the ICC (*consistency*) varies from -2.246 to $.928$, according to Table 12. These are substantial results when compared with Deniszczwicz and Kern (2013)'s results ($-.062$ to $.261$). The ICC (*consistency*) reveals the level of consistency among reviewers. For example, if Reviewer A rates 3 items with numbers 0, 3, and 6, and Reviewer B rates the same items, subsequently, with numbers 1, 4, and 7, they are very consistent because the difference of ratings between each item is the same: 1 (0-1; 3-4; 6-7). It means that even if the reviewers don't completely agree on the ratings, they might present consistency in their ratings. To exemplify, Case S19 (Table 2) present an ICC (*agreement*) of $.585$ and an ICC (*consistency*) of $.860$. The reviewers of this case agree on 58.5% of the ratings, which is already a high value, and they present consistency on 86% of their ratings.

Moreover, the reviewers present high Median values for the two types of ICC. The Median for the ICC (*agreement*) is $.348$ and for the ICC (*consistency*) is $.471$. It is much higher than Deniszczwicz and Kern (2013)'s Medians for ICC (*agreement*) of $.058$ and ICC (*consistency*) of $.097$. They are even higher than Weller's Median ICC of $.30$ (no specific ICC provided) from a study that analyzed reviews from professional researchers (Weller 2002, in Deniszczwicz & Kern, 2013).

Two possible speculations could be raised in the attempt to explain these high values: the students' high level of formal education and the positive role of the rubrics in the peer review process. As for the former, despite the fact that the students did not receive training for the peer review activity, they all have a high level of formal education and all of them work currently or worked in the past in the Education field, especially as teachers for several years.

Regarding the positive role of the rubrics, 12 students reported that the rubrics offer guidance, framework and a structure to follow. These results are from an open question of the post course survey. Although the question addresses specifically how the rubrics support their case writing, three students out of 12 highlighted the importance of the rubrics to reviewing other students' papers. Two of them reported the importance of knowing precisely what was expected from them in terms of content and how to provide feedback, as reported by a student: "I know exactly what to include in my work and what to comment on others' work" (data from post-course survey). Analogous point of view comes from this other student: "I used the written explanation of the rubrics extensively to make sure that I was being thorough and covering all the information that needed to be covered. I did the same thing when I reviewed others' works" (data from post-course survey). Another student goes further stating that the rubrics "provided a common language for giving and receiving feedback" (data from post-course survey). Perhaps well-elaborated rubrics have a strong role in both writing and peer review processes as stated by this student:

I need a rubric to give me a general idea of what is necessary and required to have a complete work. The rubrics were helpful because they had guiding

questions. The rubric was vague enough that made me feel like I had freedom to do what I thought would be best for my topic, but specific enough that I felt like I had a framework (data from post-course survey).

Despite these positive results, it is important to state that they do not determine if reviewers are right or wrong in their judgment and neither reveal the quality of the feedback. It just means that the reviewers present a high level of agreement between them and present consistency on rating when giving score for a paper. Zhang et al. (2020) investigated the reliability and validity of peer feedback across the college years of EFL learners (1st and 4th years students), and they concluded that the peer assessment had high reliability in both groups of students while it had a low validity on first year learners regarding language conventions, suggesting caution on using this type of assessment with students on their very early stage of EFL learning. Therefore, in order to scrutinize the type of descriptive feedback and to verify its quality, the following qualitative analysis was conducted.

4.2. Qualitative approach

Each criterion will be approached separately (from 1 to 6) and examples will be provided. As mentioned in the Method section, score 4 was eliminated and this level of detail was selected to have a precise perspective of what type of feedback each score demanded. It would be expected that the lower the score the better would be the feedback, as it is implied that more feedback is needed in order to achieve an optimal level. And by good quality feedback this study considers the feedback that actually articulates something that can improve the writing and that goes beyond the “cheerleader” effect (van-Haren, 2015), which just motivates the writer. This expectation was not confirmed, as forthcoming discussion. The feedback was organized into the following categories, which will be discussed in each feedback criterion: additional comments and suggestions related to the established criterion (AC/S); additional comments and suggestions unrelated to the criterion but that could improve the reviewed paper (AC/S-N); comments that have no possibilities to impact the paper (JC); and unclear comments that were impossible to understand (U).

4.2.1. Criterion 1 – The Educational Challenge

Figure 8: Description and scores for Criterion 1.

Criterion 1: The Educational Challenge

Description: Describe the background to the development of this technology. What is the educational challenge that this technology is intended to address? Reviewers: comment and suggest additional dimensions of the challenge.

Scores:

- 4: A substantial and very significant challenge
- 3: A very significant challenge.
- 2: An important challenge
- 1: A moderately important challenge
- 0: A routine challenge

Table 3 shows that 38 out of 41 descriptive feedback (92.68%) for Criterion 1 to papers developed in the Scholar course belong to the AC/S category. After analyzing all the feedback for the scores from 0 to 3, there was no difference in terms of quality or in the lengths of the feedback. In general, considering all feedback for this criterion, the shortest feedback has 12 words and the longest has 97 words.

Table 3: Quantity and Types of feedback for each score for Criterion 1

Types of Feedback	Descriptive Feedback		Total Descriptive Feedback	Total Numeric Feedback
	AC/S	JC		
Types of Score				
Score 3	19	2	21	24
Score 2	15	1	16	18
Score 1	3	-	3	3
Score 0	1	-	1	1
Total	38	3	41	46

Source: Table built by counting the results from qualitative categorization of feedback

Reviewers provide feedback for the AC/S category in various forms. Some make suggestions by using statements, as shown in the two examples below:

You have a good start to a description here. I'm wondering what the platform looks like, how students access it, and how it directly addresses a problem in the classroom. Maybe you can provide a specific problem - or several problems, since you mention that this can be utilized across the curriculum - that would help me understand why this technology is necessary. Maybe even providing some kind of vignette would be interesting and help explain the technology.

Your initial approximation of programming language to traditional language presented an unusual but pleasant background for what turned out to be an informative and interesting paper. You needed to be more explicit and cohesive in describing the challenge that the technology is designed to resolve. I was left to deduce that the challenge is that females are not as much into programming as males and that older children who have had no exposure to coding, have difficulty grasping the languages at high school.

Other reviewers make suggestions by employing questions, as can be seen in the following examples:

Is the challenge the rise in technology usage in a students non-classroom life and creating a balance? That's a great start, I like it. But, is that really the gap that game-based learning hopes to fill? Or is that just a convenient way to frame them?

There is a lot of background information here but I am not quite sure what the educational challenge is here. What hole in learning are they trying to fix?

One reviewer even provided a deeper analysis for the subject on the paper, as follows:

I think this program does more than just assess grade level. It allows schools to measure progress. Our school used this heavily to see if students were meeting growth rates. They would use this data to evaluate individual classes and see if one year has a greater growth than the next. It also allows for students to individualize their growth. A student performing below grade level could make more progress in their individual education than a student who is above grade level. The MAP program allows educators to target individual students and look at populations at large.

Concerning the feedback for the JC category, they are comprised by sentences such as “LSM is important as it is widely used in nowadays educational organizations” and “You have this pretty well thought out, though (by your own admission) the work is incomplete”, which have no content that could feed the writer with suggestions to improve the paper.

4.2.2. Criterion 2 – ‘Parse’ the Educational Technology

Figure 9: Description and scores for Criterion 2.

<p>Criterion 2: ‘Parse’ the Educational Technology</p> <p>Description: How does the technology work? Explain its structure and function from the point of view of teachers and learners. How does the way the technology works reflect its understanding of learning? Reviewers: If you believe more information is needed for a full interpretation and explanation of the technology, suggest what.</p> <p>Scores:</p> <ul style="list-style-type: none"> • 4: Excellent information coverage • 3: Good information coverage • 2: Enough information • 1: Just enough information • 0: Not enough information

Table 4 shows that 43 out of 45 descriptive feedback (95.55%) for Criterion 2 to papers developed in the Scholar course belong to the AC/S category.

Table 4: Quantity and types of feedback for each score for Criterion 2

Types of Feedback	Descriptive Feedback		Total Descriptive Feedback	Total Numeric Feedback
	AC/S	JC		
Types of Score				
Score 3	24	-	24	24
Score 2	12	1	13	14
Score 1	5	1	6	7
Total	41	2	43	45

Source: Table built by counting the results from qualitative categorization of feedback

The length of each feedback varies from 13 to 141 words. Even though the number of words used by the reviewers differed, they all offer solid comments and

suggestions that could improve the content of the paper, such as demonstrated in the following examples from reviewers in Scholar, the first being a short review and the next one a long review:

Pieces of SL are explained in a basic way that is easy to understand. Applicability to learning is still a little unclear. (22 words)

I loved the way you started this part using scientific research results. I could understand how the Geometer's sketchpad will benefit learning math. I was impressed that its product line was that diverse. I think you did really well present this part, however, if I should suggest something on your work, I would suggest you show detailed instructions at least part of them. I mean, since I have never tried this program, I wondered how it works. Now I knew the products but still I don't know how student can use it or how the program looks like. I tried the link you provided, but for some reasons, it did not work. Hence, if it is possible, I would suggest you to add more descriptions about inside of the product. (130 words)

These examples of feedback demonstrate that even short feedback can be useful to improve a paper. The first example of short feedback, for instance, using only 22 words, manages 1) to provide a general evaluation of the writing by saying that “Pieces of SL are explained in a basic way that is easy to understand” and 2) to indicate what is still unclear in this section of the writing.

4.2.3. Criterion 3 – The Underlying Learning Theory

Figure 10: Description and scores for Criterion 3.

Criterion 3: The Underlying Learning Theory

Description: What is the underlying theory of learning that this educational technology reflects? You might mention some of theorists that various members of the group introduced in the first major task in this course (cite and link to these in your references), or other learning theorists. Reviewers: comment and suggest possible additional theoretical perspectives.

Scores:

- 4: Excellent connections made with learning theories
- 3: Good connections made with learning theories
- 2: Adequate connections made with learning theories
- 1: Very little or no connections made with learning theories
- 0: Insufficient connections made with learning theories

From 55 descriptive feedback, 48 (87.27%) are related to good quality feedback, comprising the AC/S category (N=15) and the AC/S-N category (N=33), as demonstrated in Table 5. They follow the same pattern of feedback already discussed in Criteria 1 and 2. The only new element here is the type of content conveyed in the AC/S-N category.

Table 5: Quantity and types of feedback for each score for Criterion 3

Types of Feedback	Descriptive Feedback				Total Descriptive Feedback	Total Numeric Feedback
	AC/S	JC	AC/S-N	U		
Types of Score						
Score 3	9	5	10	1	25	27

Score 2	3	-	10	-	13	16
Score 1	3	1	11	-	15	12
Score 0	-	-	2	-	2	2
Total	15	6	33	1	55	57

Source: Table built by counting the results from qualitative categorization of feedback

Regarding the content of the feedback for the AC/S-N category, two of them are about citation style, such as: “The theories have been identified. But the citation of these seems not to follow the APA style. However, this could be easily improved”. The other feedback discusses the presented theories and/or deliver specific content that can improve the paper, but they fail on offering what the rubrics specifically demands: “comment and suggest possible *additional* theoretical perspectives” (my emphasis). Two instances of these occurrences:

Theories are connected and justified, but again, more specific examples or evidence would make this stronger.

I wonder if there's something else you could connect to in terms of having the simulations produced via technology--is that better or worse than a hands-on experiment? The divide between the simulations and the gamified multiplication is a bit awkward to me, because they seem very different to me. Is most of the content simulation or games? Good discussion of behaviorism and situated learning theory, but I think you could definitely expand, especially situated learning theory.

Unlike some of the “cheerleader” comments from the JC category, this is informative: “Your paper not only introduces technology but is very informative about the theory beyond the application. At every step of the paper your related your advocacy for this technology to this important educational theory”. It provides an overall evaluation of the paper, which shows to the writer the current situation of the writing. However, again, it fails in providing information that the writer could use to improve the paper.

4.2.4. Criterion 4 – The Technology in Practice

Figure 11: Description and scores for Criterion 4.

Criterion 4: The Technology in Practice

Description: Is the operation of the technology in practice adequately described in the case study? Are concrete examples provided? Do they illustrate they way the technology's underlying theory of learning translates into practice? Reviewers: What else would you like to know?

Scores:

- 0: Poor description
- 1: Adequate description
- 2: Good description
- 3: Very good description
- 4: Excellent description

Tables 6 shows that the descriptive feedback for Criterion 4 illustrates an almost perfect scenario: Forty-three from the 45 descriptive feedback (95.55%)

belong to the AC/S category. They also follow the same pattern of feedback previously discussed in Criteria 1 and 2.

Table 6: Quantity and types of feedback for each score for Criterion 4

Types of Feedback	Descriptive Feedback		Total Descriptive Feedback	Total Numeric Feedback
	AC/S	JC		
Types of Score				
Score 3	15	-	15	17
Score 2	18	-	18	22
Score 1	7	1	8	8
Score 0	3	1	4	4
Total	43	2	45	51

Source: Table built by counting the results from qualitative categorization of feedback

4.2.5. Criterion 5 – Critical Reflection

Figure 12: Description and scores for Criterion 5.

Criterion 5: Critical Reflection
Description: What are the relative strengths and weaknesses of the technology? Reviewers: provide comments and suggestions about other possible strengths and weakness that the reviewer may not have considered.
Scores:

- 4: Comprehensive analysis of strengths and weaknesses
- 3: Good analysis of strengths and weaknesses
- 2: Adequate analysis of strengths and weaknesses
- 1: Partial analysis of strengths and weaknesses
- 0: Minimal analysis of strengths and weaknesses

Answers to Criterion 5 follow the similar pattern as answers to Criterion 3 regarding good quality feedback. According to Table 7, from 43 descriptive feedback, 40 (93.02%) are related to good quality feedback, belonging to the AC/S category (N=22) and to the AC/S-N category (N=18). Additionally, they follow the identical pattern of feedback already discussed in previous Criteria in relation to forms of providing feedback (by posing questions or statements) and to the length of the feedback.

Table 7: Quantity and types of feedback for each score for Criterion 5

Types of Feedback	Descriptive Feedback				Total Descriptive Feedback	Total Numeric Feedback
	AC/S	JC	AC/S-N	U		
Types of Score						
Score 3	17	-	7	1	25	28
Score 2	1	1	9	-	11	11
Score 1	2	-	1	-	3	5
Score 0	2	-	1	1	4	4
Total	22	1	18	2	43	48

Source: Table built by counting the results from qualitative categorization of feedback

4.2.6. Criterion 6 – Conclusions and Recommendations

Figure 13: Description and scores for Criterion 6.

Criterion 6: Conclusions and Recommendations
Description: Do the conclusions and recommendations follow from the information and reasoning provided in the case study? Reviewers: comment and suggest conclusions and recommendations you might want to see added.
Scores:

- 0: Unclear conclusions and impractical recommendations
- 1: Poor conclusions and limited recommendations
- 2: Partial conclusions and incomplete recommendations
- 3: Strong conclusions and recommendations
- 4: Very clear conclusions and strong, comprehensive recommendations

As previously emphasized, the types of descriptive feedback within each category are similar for all criteria and for all scores. Table 8 displays the numbers of feedback for Criterion 6 revealing that from the 52 descriptive feedback 44 (84.61%) encompass the AC/S category, which is the good quality feedback.

Table 8: Quantity and types of feedback for each score for Criterion 6

Types of Feedback	Descriptive Feedback			Total Descriptive Feedback	Total Numeric Feedback
	AC/S	JC	U		
Types of Score					
Score 3	18	3	1	22	25
Score 2	19	2	-	21	18
Score 1	4	-	-	4	6
Score 0	3	2	-	5	6
Total	44	7	1	52	55

Source: Table built by counting the results from qualitative categorization of feedback

Additionally, descriptive feedback (Table 8) presents a higher percentage of valid good quality descriptive feedback for all criteria (categories AC/S + AC/S-N) than the results from Muck (2016) investigating Coursera. While reviews in CGScholar present a Mean of 91.4% of the valid good quality descriptive feedback, the reviews in Coursera present a Mean of 68.5%. Scholar might have more feedback because of the multimodal disposition of the platform. One possible explanation for this predisposition of reviewers offering higher quantity of descriptive feedback might be the differences in the interface between CGScholar and Coursera. As described and illustrated in the Method section, in CGScholar reviewers conduct the review online having the writing/reviewing space side-by-side with the criteria, the scores, and the boxes to type the descriptive feedback. This feature is absent in Coursera and reviewers usually have to download the paper in order to read it.

Another conclusion that can be drawn from the analyses of this section is that reviewers should be warned about the consequences or the lack of them when providing feedback. It is undeniable that people in general like receiving

compliments and approval for their achievements. However, “cheerleader” feedback *per se* as well as personal comments such as “I am glad someone is looking at including this into teacher education” (Student’s response), without any additional suggestion, is inefficient. This type of feedback does not provide information that allows possible improvement of the paper.

After all, students have to realize that this is beyond a simple task performed during a course. “The core of the peer review method for learning is the students’ change, from passive and unquestioning receptors of information, to active and critic members of a community that constructs knowledge” (Kern et al, 2007, p. 62). It refers to a change on the educational paradigm with students conquering agency and being empowered to be active producers of knowledge and agents responsible for their own learning development.

5. Concluding Remarks

The research shows that the quality of the feedback was enhanced by the resources afforded by the multimodal CGScholar platform, such as having the rubrics alongside the writing/reviewing space. According to Burbules (2010, p. 17), “virtual learning environments need to be understood not primarily in relation to technologically based ‘virtual reality’ experiences, but as immersive learning places in which creativity, problem solving, communication, collaboration, experimentation, and inquiry support a fully engaged experience”. And CGScholar supports deep and meaningful interactions. Online peer-feedback activities could be employed in all levels and modes of education to enhance writing in foreign language, as reported by (Yu & Wu, 2013).

One relevant pedagogical implication relates to teachers limited available time. As the number of students per course is increasing and the teaching hours remain the same, new modes to provide feedback are needed. At this end, the platform randomly and blindly assigns the papers to be reviewed, which enables the use of peer-feedback activities with large EFL classes either in distance or face-the-face education, enhancing the learning process, as new and “[d]ifferent skills are emerging for teaching and learning on a global scale for a global practice, including how to teach and learn in multi-time zone, multi-institutional, and multicultural settings” (Haythornthwaite, 2010, p. 42). With online peer-feedback, knowledge is co-constructed and reconceptualized among members of that learning community and with the diverse available sources of knowledge. With social knowledge technologies, such as CGScholar, the focus of the learning process is on the process (not on the product), which signals a real educational transformation.

Acknowledgements

This article reports partial results of a Doctoral Dissertation at the Graduate Program in English (PPGI) at the Federal University of Santa Catarina and supported by the Ministry of Education of Brazil, through Capes Foundation

(Coordination for the Improvement of Higher Education Personnel), and by the College of Education of the University of Illinois at Urbana-Champaign (UIUC). Our appreciation to Professor Viviane Heberle, advisor to the first author, and to Kathleen Santa Ana, from the Applied Technologies for Learning in the Arts & Sciences - UIUC, for her technical support.

References

- Al Abri, A.; Al Baimani, S.; Bahlani, S. (2021). The Role of Web-based Peer Feedback in Advancing EFL Essay Writing. *Computer-Assisted Language Learning Electronic Journal (CALL-EJ)*, 22(1), 374-390. Available at <http://callej.org/journal/22-1/Al%20Abri-Al%20Baimani-Al%20Bahlani2021.pdf>.
- Al Qunayeer, H. S. (2020). Exploring EFL learners' online participation in online peer writing discussions through a Facebook group. *Journal of Information Technology Education: Research*, 19, 671-692. <https://doi.org/10.28945/4621>.
- Alharbi, M. A. (2019). Exploring the potential of google doc in facilitating innovative teaching and learning practices in an EFL writing course. *Innovation in Language Learning and Teaching*. 14(3), p. 227-242. <https://doi.org/10.1080/17501229.2019.1572157>.
- Burbules, N. C. (2010). Meaning of 'Ubiquitous Learning'. In: Ubiquitous Learning. Eds. Bill Cope and Mary Kalantzis. University of Illinois Press: Urbana and Chicago.
- Butin, D. W. (2012). What MIT should have done. *eLearn Magazine*.
- Cheng, K.-H., Liang, J.-C., & Tsai, C.-C. (2015). Examining the role of feedback messages in undergraduate students' writing performance during an online peer assessment activity. *Internet and Higher Education*, 25, 7.
- Cope, B.; Kalantzis, M. (2013). Towards a New Learning: the Scholar social knowledge workplace, in theory and practice. *E-Learning and Digital Media*, 10(4), 25.
- Dambros, M. T. (2020). Ferramenta de tecnologia educacional WhatsApp na construção do conhecimento. Professional Masters Course Report. Mestrado Profissional em Práticas Transculturais. Centro Universitário Facvest, Lages, SC. (Unpublished).
- Deniszczwicz, M., & Kern, V. M. (2013). Indicadores de gestão na revisão por pares: confiabilidade da revisão recíproca anônima de propostas de mestrado. *Liinc em Revista*, 9(1), 13.
- Dornyei, Z. (2007a) *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- Dornyei, Z. (2007b). Qualitative, quantitative, and mixed methods research *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- Elfiza, R.; Reszki, A. S.; Nopita, D. (2021). The Effect of Online Written Feedback through Social Media on Students' Writing. *J-SHMIC: Journal of English for Academic*, 8 (1). [https://doi.org/10.25299/jshmic.2021.vol8\(1\).6034](https://doi.org/10.25299/jshmic.2021.vol8(1).6034).
- Haythornthwaite, C. (2010). Participatory Transformations. In: Ubiquitous Learning. Eds. Bill Cope and Mary Kalantzis. University of Illinois Press: Urbana and Chicago.
- Heberle, V. M., & Abreu, B. B. (2011). Investigating multimodality: an analysis of students' diary journals. *Revista Linguagem Em Foco*, 3(5), 77-96. Retrieved from: <https://revistas.uece.br/index.php/linguagememfoco/article/view/1850>

- Gikandi, J. W.; Morrow, D.; Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, 57(4), 19.
- Jeong, K. (2016) A Study on the Integration of Google Docs as a Web-based Collaborative Learning Platform in EFL Writing Instruction. *Indian Journal of Science and Technology*, 9(39). DOI: 10.17485/ijst/2016/v9i39/103239.
- Kalantzis, M.; Cope, B. (2012). *New Learning: Elements of a science of education*. Cambridge: University of Cambridge Press.
- Kern, V. M., Pacheco, R. C. d. S., Saraiva, L. M., & Pernigotti, J. M. (2007). Peer Review in Computer Science: Toward a Regular, Large Scale Educational Approach. In F. M. Mendes & F. V. Brasileiro (Eds.), *Advances in Computer-Supported Collaborative Learning*. Hershey, USA: Idea Group Inc.
- Kern, V. M., Pernigotti, J. M., Calegari, M. M., & Bento, M. (2002). *Peer review in engineering education: Speeding up learning, looking for a paradigm shift*. Paper presented at the INTERTECH '2002 – International Conference on Engineering and Technology Education, Santos - Brazil.
- Kern, V. M., Saraiva, L. M., & Pacheco, R. C. d. S. (2003). Peer Review in Education: Promoting Collaboration, Written Expression, Critical Thinking, and Professional Responsibility. *Education and Information Technologies*, 8(1), 10.
- Kollar, I., & Fischer, F. (2010). Peer assessment as collaborative learning: a cognitive perspective. *Learning and Instruction*, 20(4), 5.
- Locke, L. F., Silverman, S. J.; Spirduso, W. W. (1998) *Reading and understanding research*. Thousand Oaks, CA: Sage Publications.
- Muck, K. (2015). The Role of Recursive Feedback: A Case Study of e-Learning in Emergency Operations. *The International Journal of Adult, Community, and Professional Learning and Instruction*, 23(1), 25.
- Muck, K. E. & Sadki, R. (2015). Learning in Emergency Operations: A Case Study of a Cross-Sector Distance Learning Course Organized by the International Federation of Red Cross and Red Crescent Societies (IFRC). LSi. DOI: 10.13140/RG.2.2.26960.94721.
- Noroozi, O., Biemans, H., & Mulder, M. (2016). Relations between scripted online peer feedback processes and quality of written argumentative essay. *Internet and Higher Education*, 31, 12.
- Paker, T.; Dogan, N. (2021). Using Edmodo as a Social Platform to teach EFL Writing for Preparatory School Students. *Journal of Language Education and Research* 7(1). DOI: 10.31464/jlere.850794.
- Pol, J. v. d., Berg, B. A. M. v. d., Admiraal, W. F., Simons, P. R. J. (2008). The nature, reception, and use of online peer feedback in higher education. *Computers & Education*, 51, 14.
- Seyyedrezaie, Z. S.; Ghonsooly, B.; Shahriari, H.; Fatemi, A. H. (2016). A mixed methods analysis of the effect of google docs environment on efl learners' writing performance and causal attributions for success and failure urkish. *Online Journal of Distance Education-TOJDE*. 17(3). Available at <https://dergipark.org.tr/en/download/article-file/222623>.
- Slavkov, N. (2015). Sociocultural Theory, the L2 Writing Process, and Google Drive: Strange Bedfellows? *Tesl Canada Journal/Revue Tesl Du Canada*. 32(2). Available at <https://files.eric.ed.gov/fulltext/EJ1083966.pdf>.
- van_Haren, R. (2015). Educational Blog Restricted to a Class. CGScholar. - Private Access.

- Yang, Y.-F. (2016). Transforming and constructing academic knowledge through online peer feedback in summary writing. *Computer Assisted Language Learning*, 29(4), 20.
- Yu, F.-Y., & Wu, C.-P. (2013). Predictive Effects of Online Peer Feedback Types on Performance Quality. *Educational Technology & Society*, 16(1), 10.
- Zhang, F.; Schunn, C.; Wentao, L.; Long, M. (2020). Changes in the reliability and validity of peer assessment across the college years. *Assessment & Evaluation in Higher Education*, 45(8), 1073–1087. DOI: <https://doi.org/10.1080/02602938.2020.1724260>.

Recebido em: 15/04/2021

Aceito em: 03/05/2021