Testing the Efficacy and Student Acceptance of a Peer-Review Writing Program in an Online Course

Andrew M. Bouwma

Oregon State University

Oregon State University

Abstract

Writing within Discipline is a pedagogical model that aims to engage students in active learning and to develop critical thinking and writing skills within the norms of a particular discipline. However, lack of faculty time for grading writing products impedes its broader implementation. Web based peer review programs have been developed that allow for writing assignments in large classes without overburdening instructors. Yet unknowns and barriers remain regarding successful implementation of these programs, notably efficacy in online courses and students' comfort with and acceptance of the programs. The current study investigated learning gains and student acceptance of the peer review program, Peerceptiv, when used as a tool for providing feedback on term papers in an online biology course. Students showed high acceptance of the Peerceptiv assignments, and demonstrated significant learning gains.

Introduction

Writing in the Disciplines (WID) is a pedagogical model that aims to engage students in active learning and to develop their critical thinking and writing skills within the norms of a particular discipline (Oschner and Fowler, 2004). WID is not a new model, but lack of faculty time for grading writing products impedes its broader implementation (Persky et al., 2003). Oregon State University (OSU) has embraced WID with its Writing Intensive Curriculum (WIC), which requires students to complete at least one WIC course within their major before they graduate. OSU WIC course sections are limited to 20 student enrollments in order to make the required 5,000 total words of writing manageable for the instructor. Other OSU [Baccalaureate Core Courses](#), with higher enrollments, also have WID-inspired writing requirements, such as the Synthesis category, which requires a 1250-word term paper. Providing meaningful feedback to students and the opportunity for revision can be difficult for instructors in Synthesis category classes with 60+ students.

Within the past decade, web-based peer review software programs have been developed that allow for frequent writing assignments to be given to large classes without overburdening instructors or teaching assistants (Cho and Schunn, 2007; Clase et al., 2010). One such online review program, Scaffolded Writing and Rewriting within the Disciplines (SWoRD; now known as Peerceptiv) (Cho and Schunn, 2007), has been shown to be an effective tool at managing writing assignments and providing students with feedback of high quality, useful for subsequent revision/improvement of writing assignments (Patchan et al., 2011).

Peerceptiv is a web-based program that provides double-blinded peer review for writing assignments, and is modeled after the process of academic writing (Cho and Schunn, 2007). For the first stage of an assignment using Peerceptiv, students compose a writing assignment and submit it via the web. For the second step of the assignment, each student is anonymously assigned as reviewer of three (or more) of their peers' papers. The students then review each paper using highly detailed 10-point rubrics, also providing required comments to justify their score choice, and suggestions for improvement and praise for things well done. The third step of the assignment is "feedback," where students are required to review their scores, read the anonymous comments left by their reviewers, and rate the comments on their degree of helpfulness. Upon completion of the assignment, students may appeal directly to the instructor any peer scores that they believe to be inaccurate, and the instructor is able to override any scores they determine to be inaccurate.

In Peerceptiv assignments, students are scored on three categories: Writing, Reviewing, and Task scores. The writing score is based on a weighted average that comes from their peers' scores on the 10-point scale. The reviewing score is based on the proximity of the student's reviews relative to the mean review given to papers by their peers, and the helpfulness rating given to them by their peers for their reviews. The task score awards points for completion of all required review and feedback

tasks on time. Thus, students are given strong incentive to provide accurate, timely, and detailed feedback to their peers. The weight given to each of the three categories on which students are graded is adjustable based on instructor preference. Thus, students ultimately receive feedback from the ratings, comments provided by their peers, feedback on the quality of their reviews from their peers, and feedback from their instructor in the form of the detailed grading rubrics used for the evaluation. Peerceptiv allows for multiple drafts of a paper to be assigned and for each draft to serve as one assignment affording peer and instructor review. This allows for the complete cycle of academic writing, with submission, peer review, revision and resubmission. An example rubric is shown in Appendix A.

Studies with students in on-campus courses have shown that online peer review programs that utilize peer grading, such as Peerceptiv, are effective in providing students with useful feedback (Patchen et al., 2009) and improving writing assignment quality (Patchan et al., 2011). To the best of my knowledge, no formal, published studies exist regarding efficacy of such programs for learning, and writing development specifically, for students participating in fully online courses (See Coates, 2017, for an unpublished case study on Peerceptiv in an online entomology course). The relatively small body of literature regarding the efficacy of peer review programs writ large points to a problem of student resistance to such systems, that can ultimately prevent faculty from utilizing them (Kaufman and Schunn, 2011; Keeny-Kennicutt, 2008). Moreover, students sometimes question the value of obtaining feedback from peers who are learning course material at the same time as they are, and are more likely to have positive perceptions of such systems if the instructor is involved in assigning grades from peer review (Kaufman and Schunn, 2011).

The current study addressed the following two research questions: 1) To what extent will students in an online science course show writing gains in Peerceptiv assignments? 2) If students are given evidence from the literature that demonstrates the value of peer feedback for their writing, and are reminded of the robust course policies allowing for appeal of scores, will their perceptions of Peerceptiv assignments be impacted?

Methods

This study was conducted during summer and fall 2018 quarters with two sequential sections of the online Zoology course Z349: Biodiversity: Causes, Consequences, and Conservation. This course requires a 1,250 word term paper. The course enrollment (approximately 60 students per quarter) precludes the instructor from giving the students traditional, timely feedback, in the form of comments on drafts and opportunities for revision before they turn in their final term paper. The lead investigator used the Peerceptiv program in order to provide the students with incremental, peer-generated feedback on their term papers. Student term papers were produced in three parts with each part a separate Peerceptiv assignment. Students received feedback from the Peerceptiv peer reviewers on their drafts, and then synthesized the three parts (1, 2, and 3) into a final paper, which was graded by the instructor.

The summer 2018 course was taught by the lead investigator, while the fall 2018 course was taught by a colleague of the lead investigator. To reduce researcher bias, in the summer course, student participants were recruited via an email sent by a graduate student assistant, and in the fall course, students were recruited via an email sent by the lead investigator. For summer 2018 the rate of study participation was 37.5% (15/40) and for fall 2018 it was 31.4% (16/51).

Each term, the student participants were randomly assigned to two equal-sized groups: Group One received the intervention, while Group Two was the control group. All students in the study were assigned codes and all data were de-identified, so that no identifiable student records were maintained. Students were offered a total of 20 points of course extra credit (out of 600 total course points) for their full participation in the study. Because some students dropped the course, stopped participating in the study, or did not

complete all of the assignments, the sample sizes for each group were not equal at the end of the study (see Table 1).

Table 1: Study sample sizes

Group One		
	Assignments	Final Survey
Summer 2018	6	7
Fall 2018	5	6
Subtotals	11	13
Group Two		
	Assignments	Final Survey
Summer 2018	8	8
Fall 2018	10	10
Subtotals	18	18
Totals	**29**	**31**

Research Question 1: Writing gains

Research participants in both Group One and Two, like all others enrolled in the course, produced three writing products using Peerceptiv. The investigators rated assignments #1 and #3 for all study participants from the courses in summer 2018 (n = 14) and fall 2018 (n = 15) using the peer review rubric that was part of the Peerceptiv assignments. Assignments #1 and #3 were selected for comparison in order to increase the likelihood that learning gains could be observed, since this allowed students two assignments (#1 and #2) with peer review in order to make learning gains before their final assessment. In addition, the students rated the same papers as part of the review portion of the Peerceptiv assignments. Assignment #1 and assignment #3 median rubric scores, compiled by both the investigators and the students, were compared via paired Wilcoxon sign rank tests, which allowed for the comparison of the ordinal (Likert) rubric scores.

Research Question 2: Perceptions of Peerceptiv

Research participants who were randomly assigned to Group One (intervention) at the beginning of the study were given two short video interventions, one that described the research behind peer review and the Peerceptiv program itself, and a second that described the Z349 peer review appeal process if students wished to challenge a perceived inaccurate Peerceptiv peer review. Students in Group One were quizzed on the content of the videos (for 12 points of extra credit) to ensure that they watched them. Students in Group Two, the control group, were shown two different videos of comparable length about ecological economics, rather than Peerceptiv, and also took a quiz worth the same amount of extra credit (12 points). The quizzes in the study were administered in an external online survey platform so that the instructor was not aware of participation of individual students until after the end of the term. At the end of the course, each group was surveyed on their perception of the fairness of Peerceptiv peer review assignments, as well as of their self-reported learning gains in regards to course content and writing skills using a 6-point Likert scale for agreement (1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = somewhat agree, 5 = agree, 6 = strongly agree). Participants earned 8 additional points of extra credit for completing the survey. Survey questions are shown in Appendix B. Potential students who did not wish to participate in the study but who wished to earn extra credit were offered the option to watch the same videos as Group Two and to take the same quiz for an equal amount of total extra credit (20 points). The Likert scale responses for each question on the final online survey were compared between Groups One (n = 13) and Two (n =18) using a Mann-Whitney U Test (because of unequal sample sizes).

Results

Research Question 1: Writing gains

The results of the Wilcoxen sign rank tests showed significant learning gains between Peerceptiv assignments #1 and #3, both when papers were scored by students and when scored by the investigators. When comparing student ratings for Peerceptiv Assignments #1 and #3, the instructor ratings and the student ratings agreed that students improved on their thesis statements and

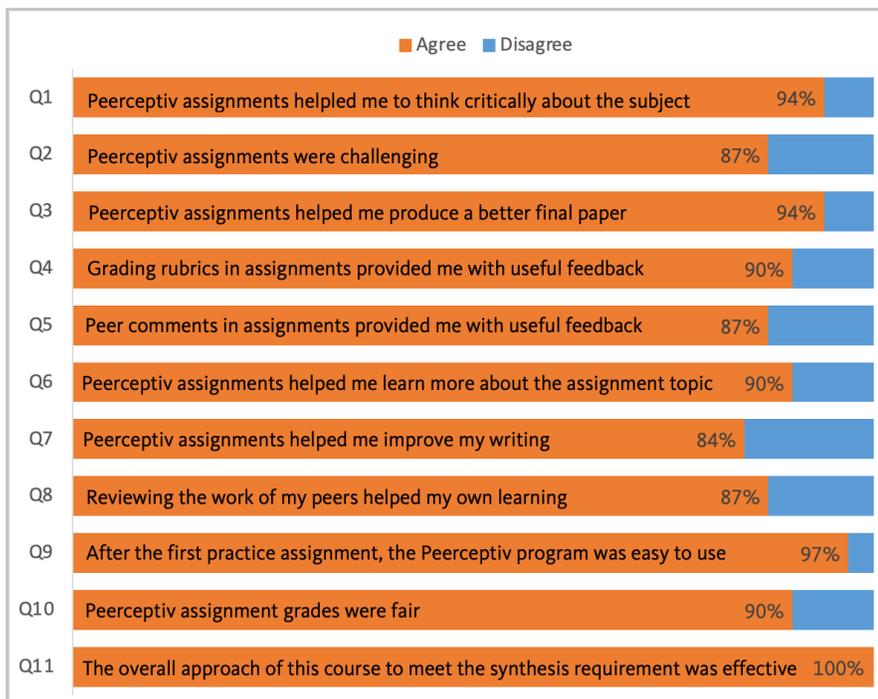essay structure (see Table 2). In addition, student raters saw improvement in the in-text citations over time, while the investigators did not. All other comparisons, for topic sentences, proofreading, and format of references, were non-significant ($p > .05$).

Table 2: Changes in performance assessed by students & investigators

N = 29	Students		Investigator	
	Reviewer type			
Rubric category	z-statistic	p-value	z-statistic	p-value
Thesis statement	-2.23	**.026**	-2.61	**.009**
Topic sentences	-1.76	.078	-0.69	.490
Proofreading	-0.08	.936	0.02	.984
Essay structure	-2.21	**.027**	-2.99	**.003**
Reference formatting	-0.24	.810	-0.24	.810
In-text citations	-2.37	**.018**	0.33	.741

Research Question 2: Perceptions of Peerceptiv

The results of the final online survey showed that Groups One and Two both showed high acceptance of Peerceptiv peer review assignments (see Figure 1). Overall agreement percentages ranged from 84 to 100% for the 11 survey questions (see Figure 1, Appendix B).

Agree / Disagree

- Q1 — Peerceptiv assignments helped me to think critically about the subject — 94%
- Q2 — Peerceptiv assignments were challenging — 87%
- Q3 — Peerceptiv assignments helped me produce a better final paper — 94%
- Q4 — Grading rubrics in assignments provided me with useful feedback — 90%
- Q5 — Peer comments in assignments provided me with useful feedback — 87%
- Q6 — Peerceptiv assignments helped me learn more about the assignment topic — 90%
- Q7 — Peerceptiv assignments helped me improve my writing — 84%
- Q8 — Reviewing the work of my peers helped my own learning — 87%
- Q9 — After the first practice assignment, the Peerceptiv program was easy to use — 97%
- Q10 — Peerceptiv assignment grades were fair — 90%
- Q11 — The overall approach of this course to meet the synthesis requirement was effective — 100%

*Note: The response scales were combined as follows: Agree = Strongly Agree, Agree, and Somewhat Agree; Disagree = Strongly disagree, Disagree, and Somewhat Disagree.

Figure 1. Student responses to survey questions regarding acceptance of Peerceptiv assignments (Groups One and Two combined).

Ninety percent of the students surveyed agreed that the Peerceptiv assignments were fair, and 94% agreed that the Peerceptiv assignments made them think critically about the material and helped them to produce a better final paper. Almost all, (97%) of the students found the program easy to use. Notably, all of the students surveyed agreed that the approach with Peerceptiv assignments used in the course was an effective way to meet the term paper requirement.

The results of the Mann-Whitney U Tests comparing the survey responses between the two treatment groups were not statistically significant (see Table 3). Thus, while student acceptance was high across both groups, the intervention did not make a difference in student acceptance of Peerceptiv.

Table 3: Student impressions of Peerceptiv by treatment group

Survey question	Group One (n = 13)		Group Two (n = 18)		Mann-Whitney U p-value
	Mean	Median	Mean	Median	
Q1	4.77	5	4.89	5	.62
Q2	4.62	5	4.72	5	.75
Q3	4.77	5	5.11	5	.57
Q4	4.69	5	4.94	5	.73
Q5	4.62	5	4.67	5	.95
Q6	4.92	5	4.83	5	.86
Q7	4.39	5	4.56	5	.83
Q8	5.08	5	4.89	5	.49
Q9	4.77	5	5.44	5	.11
Q10	5.08	5	4.72	5	.62
Q11	5.31	5	5.17	5	.60

Ratings based on a 1 − 6 Likert scale

Discussion

An important principle of scientific teaching is that teaching methods should periodically be rigorously assessed based on data in order to determine if they are helping students meet learning objectives (Handelsman et. al 2004). Documenting learning gains over sequential assignments that align with course objectives is a way to show that teaching methods are effective (Handelsman et. al 2007).

Several significant learning gains were observed over the course of sequential Peerceptiv assignments in the current study. The students and the investigators agreed that papers improved their thesis statements and the overall structure of their essays over time. Since the first assignment

required an expository thesis statement, while the third assignment required an argumentative thesis, we cannot rule out that students simply had an easier time writing argumentative versus expository theses. However, this study clearly showed that students did improve the overall structure of their essays over the course of the Peerceptiv peer review process. In addition, student reviewers observed that the in-text citations improved between the assignments, while the investigators did not. The lack of concordance between the two sets of evaluators suggests that the assignment author should revise the rubric and/or improve the assignment instructions for in-text citations. The lack of

improvement in topic sentences, proofreading, and format of references also indicates that the assignment author should review the assignment instructions and/or peer rubrics in order to increase the likelihood that students will achieve learning gains in these areas. Such iterative improvement of assignments based on data is an important goal of Scientific Teaching (Handelsman et al., 2007).

Given previous studies that have documented student resistance to peer review systems, (Kaufman and Schunn, 2011; Keeny-Kennicutt, 2008), the authors expected to observe more negative perceptions of the Peerceptiv assignments. Rather, in the current study, students were highly accepting of Peerceptiv assignments and peer grading, regardless of treatment. This result suggests that, from the point of view of students, Peerceptiv assignments utilized in the study course are an effective means to engage students in WID in an online science course. Indeed, the peer review assignment model used in Z349 could be used a model for similar courses in which a term paper is required. In addition, in order to further explore the utility of peer review systems for teaching biology, future studies could examine the efficacy of using Peerceptiv writing assignments to teach specific biology concepts in lieu of traditional reading assignments and quizzes.

References

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the disciplines: A web-based reciprocal peer review system. *Computers & Education, 48(*3), 409-426.

Clase, K. L., Gundlach, E., & Pelaez, N. J. (2010). Calibrated Peer Review for Computer-Assisted Learning of Biological Research Competencies. *Biochemistry and Molecular Biology Education, 38*(5), 290-295. doi:10.1002/bmb.20415

Coates, C. J. (2017). Case Study: Peerceptiv - Peer Assessment in the Sciences. Retrieved from https://peerceptiv.com/resources/peerceptiv-case-study-peer-assessment-in-the-sciences/

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., . . . Wood, W. B. (2004). Scientific teaching. *Science, 304*(5670), 521-522. doi:10.1126/science.1096022

Handelsman, J., Miller, S., & Pfund, C. (2007). Scientific Teaching. New York: W.H. Freeman.

Kaufman, J., & Schunn, C. (2011). Students' perceptions about peer assessment for writing: their origin and impact on revision work. *Instructional Science, 39*(3), 387-406. doi:10.1007/s11251-010-9133-6

Keeny-Kennicutt, W., Baris Gunersel, A., & Simpson, N. (2008). Overcoming student resistance to a teaching innovation. *International Journal for the Scholarship of Teaching and Learning, 2(1)*, 5. doi.org/10.20429/ijsotl.2008.020105

Ochsner, R., & Fowler, J. (2004). Playing devil's advocate: Evaluating the literature of the WAC/WID movement. *Review of Educational Research, 74*(2), 117-140. doi:10.3102/00346543074002117

Panther Learning. (n.d.). Peerceptiv – How it Works. Retrieved from https://peerceptiv.com/how-it-works/

Patchen, M. M., Charney, D., & Schunn, C. D. (2009). A validation study of students' end comments: Comparing comments by students, a writing instructor, and a content instructor. *Journal of Writing Research, 1*(2), 124-152.

Patchen, M. M., Schunn, C. D. & Clark, R. J. (2011). Writing in natural sciences: Understanding the effects of different types of reviewers on the writing process. *Journal of Writing Research, 2*(3), 365-393.

Persky, H. R., Daane, M. C., & Jin, Y. (2003). The nation's report card: Writing 2002 (NCES 2003-529). US Department of Education. Institute of Education Sciences. National Center for Education Statistics. Washington D.C. Government Printing Office.

Appendix A: Peerceptiv Rubric

Mechanics: Thesis Statement	Please rate the thesis statement according to the rubric
7 - Distinguished	Meets ALL FIVE of the following criteria: 1. Thesis statement is clear 2. expository 3. the last sentence of the first paragraph 4. bolded 5. and all information supports the thesis
5 - Proficient	Meets THREE OF THE FIVE criteria
3 - Emerging	Meets TWO OF THE FIVE criteria
1 - Not evident	No identifiable thesis statement in the introductory paragraph

Mechanics: Topic Sentences	Please rate the topic sentences according to the below rubric
7 - Distinguished	Topic sentences clearly introduce the main idea of the paragraph
5 - Proficient	Topic sentences mostly introduce the main idea of the paragraph
3 - Emerging	Topic sentences somewhat introduce the main idea of the paragraph
1 - Not evident	No consistently identifiable topic sentences that indicate the main idea of the paragraph

Mechanics: Proofreading	Please rate the degree to which the essay was reviewed for typos, misspellings, and grammatical mistakes
7 - Distinguished	The essay has no evident typos, misspellings and grammatical mistakes
5 - Proficient	The essay has at least THREE evident typos, misspellings or grammatical mistakes
3 - Emerging	The essay has at least SIX evident typos, misspellings or grammatical mistakes
1 - Not evident	The essay has at least TEN OR MORE evident typos, misspellings or grammatical mistakes

Mechanics: Essay Structure	Please rate the essay with the below rubric based on its overall structure
7 - Distinguished	The essay has an introductory paragraph that introduces the topics that are discussed, body paragraphs that explain each topic, and a concluding paragraph that sums the essay
5 - Proficient	The essay has an introductory paragraph that introduces most of the topics that are discussed, body paragraphs that explain each topic, and a concluding paragraph that somewhat sums the essay

3 - Emerging	The essay has an introductory paragraph that introduces some of the topics that are discussed, body paragraphs that explain some topics, and a concluding paragraph that sums part of the essay
1 - Not evident	The essay is not clearly organized into introductory, body, and concluding paragraphs

Mechanics: Format of references	Please rate the references according to the below rubric
7 - Distinguished	References page contains at least 5 appropriate APA formatted citations, and at least two of them come from a peer-reviewed source
5 - Proficient	References page contains less than 5 appropriate APA formatted citations with minor errors present, and at least two of them come from a peer-reviewed source
3 - Emerging	References page contains less than 5 appropriate APA formatted citations with minor errors present, and/or fewer than two of them come from a peer-reviewed source
1 - Not evident	References page missing, major problems with formatting, or not peer-reviewed references

Appendix B: Final Survey Questions

Please indicate the degree to which you agree or disagree with the following statements. 1 = strongly disagree, 2 = disagree, 3 = somewhat disagree, 4 = somewhat agree, 5 = agree, 6 = strongly agree	
Q1	Peerceptiv assignments helped me to think critically about the subject.
Q2	Peerceptiv assignments were challenging.
Q3	Peerceptiv assignments helped me produce a better final paper.
Q4	Grading rubrics in Peerceptiv assignments provided me with useful feedback that I could use to revise my paper.
Q5	Peer comments in Peerceptiv assignments provided me with useful feedback that I could use to revise my paper.
Q6	Peerceptiv assignments helped me learn more about the assignment topic.
Q7	Peerceptiv assignments helped me improve my writing.
Q8	Reviewing the work of my peers helped my own learning.
Q9	After the first practice assignment, the Peerceptiv program was easy to use.
Q10	Peerceptiv assignment grades were fair.
Q11	The overall approach of this course to meet the Bacc Core Synthesis requirement of a 1250 word paper was effective.

About the Research Unit at Oregon State Ecampus

Vision

The Ecampus Research Unit strives to be leaders in the field of online higher education research through contributing new knowledge to the field, advancing research literacy, building researcher communities and guiding national conversations around actionable research in online teaching and learning.

Mission

The Ecampus Research Unit responds to and forecasts the needs and challenges of the online education field through conducting original research; fostering strategic collaborations; and creating evidence-based resources and tools that contribute to effective online teaching, learning and program administration.

Contact

Naomi R. Aguiar, Ph.D.
Assistant Director of Research
Oregon State Ecampus
541-737-4629
naomi.aguiar@oregonstate.edu

Creative Commons License

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Suggested Citation

Bouwma, A. M. (2021). Testing the Efficacy and Student Acceptance of a Peer-Review Writing Program in an Online Course. *White Paper*. Corvallis, OR: Oregon State University Ecampus Research Unit.