How Timing and Authority in Peer Review Impact STEM Students: A Comparative Assessment of Writing and Critical Thinking in Kinesiology Courses

Tereza Joy Kramer, California Northstate University
Joe Zeccardi, Saint Mary's College of California
Chi-An W. Emhoff, Saint Mary's College of California
Claire Williams, Saint Mary's College of California
Robin J. Dunn, Saint Mary's College of California
Joshua Rose, College of Alameda

Abstract: This comparative, mixed-methods study illustrates the impact of weekly facilitated peer review ("Writing Circles") in STEM courses across time: 1) in a lower-division course, Circles improve all learning outcomes for writing and critical thinking, and most significantly, writing; 2) in an upper-division course, Circles are most effective at improving learning outcomes for critical thinking; 3) when comparing scores in the lower- and upper-division courses, we see that critical thinking improves significantly from second to fourth year; 4) finally, we see that upper-division students grant their peers more disciplinary authority during the Circles peer review.

Thinking about pedagogy for writing in the disciplines (WID) with respect to time, we distinguish three kinds of questions. First, how much time should students spend on learning to write in the discipline (Deans, 2017; Graham, 1992; Kramer et al., 2019)? Second, given some quantity of time, how should students spend it (Armstrong & Paulson, 2008; Bruffee, 1984; Gere, 1987; Kramer et al., 2019); that is, what should they be doing? Bruffee (1984), Gere (1987), Kramer et al. (2019), and others recommend collaborative learning, specifically peer review. Brieger & Bromley (2014), reporting on their model for facilitating peer review, summarize many of the advantages of incorporating peer review in undergraduate science, technology, engineering, and math (STEM) writing assignments. These include more critical thinking "through improved conceptualization, synthesis, evaluation, and application of new information" (p. 2). Indeed, as vehicles for collaborative learning, peer review groups are one way to decentralize authoritative barriers to entry (Keating, 2019), which can frustrate the efforts of STEM students seeking to join disciplinary discourse communities.

Proceeding from this premise, then, the third question arises: the timing of time, or more specifically, at what points in a college education should a student spend time learning and practicing peer review (Beaufort, 2007; Gere, 2019; Keating, 2019)? This third question includes multiple considerations. Are learning outcomes for writing and critical thinking differently impacted by peer review at different stages of a college career? And how do psycho-socio-epistemic dimensions of peer review manifest at different

Across the Disciplines wac.colostate.edu/atd
A Journal of Language, Learning and Academic Writing ISSN 554-8244
https://doi.org/10.37514/ATD-J.2022.18.3-4.06

stages; specifically, how do students perceive each other's authority in the discipline, and how does that perception impact their learning?

Given the ubiquity and necessity of peer review in the sciences generally, an observer might expect to find it similarly situated in STEM classrooms devoted to preparing students for writing and research in the sciences. Instead, as is the case in many disciplines but perhaps in STEM more acutely, the learning and practice of content-area knowledge takes precedence over the learning and practice of peer review. Furthermore, content-area experts rarely are also experts in teaching the writing process, including peer review. Therefore, when and how to effectively develop students' ability to write is a central concern for educators invested in preparing students for successful careers in the sciences. Kramer et al. (2019) note that "student peer review can be unhelpful or even backfire…We ourselves have witnessed un- or under-structured peer review resulting in students giving each other too little, incorrect, or otherwise unhelpful advice" (p. 29).

While forms of student peer review have long been adopted and researched in the field of writing studies (e.g., Gere, 1987; DiPardo & Warshauer Freedman, 1988; Cho & Schunn, 2007), there continues to be a need to better understand its benefits and how they are realized over time in discipline-specific courses. Kimberly Baker (2016), studying peer review within individual sociology courses, notes that "while numerous studies have investigated peer review, these studies have focused primarily on outcomes rather than process" (p. 4). In STEM disciplines, Julia Reynolds and colleagues (2012) say there is a general lack of writing-to-learn strategies, which employ writing as means for students to gain understanding of science concepts and methods. Undertaking a National Science Foundation-funded study and drawing on the expertise of 12 STEM research and education experts, Reynolds et al. built on a previous review (Rivard, 1994) to identify empirically validated writing-to-learn practices in STEM (Gerdeman et al., 2007; Pelaez, 2002; Walvoord et al., 2008). More recent research on the adoption of peer review in STEM courses suggests that it can be beneficial (Finkenstaedt-Quinn et al., 2019; Finkenstaedt-Quinn et al., 2021; Reddy et al., 2021; Yalch et al., 2019).

Important considerations revolve around the guiding and timing of student peer review. Gere (2019) and Keating (2019) show that students understand, experience, and respond to peer review in different ways at different stages of their education. In the early stages of their studies, students regard neither themselves nor their peers as legitimate, authoritative sources of knowledge. As a result, they tend to resist peer review, experience it negatively, and describe it as a frustrating, inauthentic waste of time. Conversely, students in the later stages of their education begin to understand themselves and their peers as authorized sources of knowledge, and are therefore more likely to value the feedback they receive in peer review. Given Keating's (2019) account of students' evolving understanding and experience of peer review, it remains to be discerned not just how and to what extent these different perspectives manifest in student writing but also when.

Whereas Gere's (2019) and Keating's (2019) studies are broad surveys of students writing in 47 different majors, our research focused in a detailed way on a single discipline, kinesiology. Further, whereas our previous work, Kramer et al. 2019, examined the impacts of iterated, structured, facilitated peer review groups ("Writing Circles" or "Circles") at a single developmental point, this study compares the impact of Circles on student writing and critical thinking at different stages in their major. Accordingly, this study analyzes student writing from a lower-division WID course in Research Methods & Writing (RMW) typically taken in the second year and an upper-division lab course in Exercise Physiology (EP) typically taken during the fourth year. We employed mixed methods, using quantitative analyses to compare disciplinary writing produced with Circles at both stages to writing produced without Circles, and qualitative analyses of reflective writing to compare lower-division Kinesiology students' understanding and experience of Circles to that of upper-division Kinesiology students. Our quantitative research questions focused on the ways in which these differences manifest in student writing, specifically: How does peer review impact students' writing at different stages of their disciplinary and psycho-socio-epistemic

development? And how do the impacts of peer review practiced among novice groups differ from those with more disciplinary expertise? Our qualitative research questions asked how and to what extent students' understanding and experience of Circles changed as they progressed through the major. Below, we report the quantitative results first, followed by the qualitative results. Our data indicate that the benefits of peer review advance in parallel with evolutions in students' understanding and experience of peer review, and that peer review manifests in different ways at different times in students' cognitive and disciplinary development.

Context on Kinesiology and Circles

Kinesiology at Saint Mary's College of California is an interdisciplinary major. In addition to courses taken through the Department of Kinesiology, students take courses in the natural sciences, social sciences, and business, based upon their chosen area of emphasis: Exercise Science, Health Promotion, or Sport & Recreation Management. As a consequence, students majoring in kinesiology exhibit a broad spectrum of writing abilities, styles, processes, and backgrounds that might be found in other majors across campus.

The lower-division course Research Methods & Writing in Kinesiology (RMW) provides students the opportunity to consider fundamental research questions in kinesiology, read and interpret research articles, and explore issues related to evaluation and measurement techniques. As their final project, students write a research proposal consisting of an extensive literature review and a detailed proposal for an experiment (Kramer et al., 2019). The RMW course immerses students in empirical studies in order to familiarize them with the research methodology used in kinesiology. The course includes an overview of various types of discipline-specific writing and an introduction to research, measurement, and evaluation within the kinesiology discipline.

The upper-division course Exercise Physiology (EP), which has RMW as one of its prerequisites, requires students to enroll concurrently in a lecture and lab to learn theoretical concepts and gain proficiencies in applied laboratory skills, respectively. Throughout the EP course, students write four lab reports, each progressing in complexity and breadth of content, based on hands-on experiments conducted in labs, incorporating theory learned in the lecture portion of the course. These lab reports are intended to provide a deeper inquiry and discussion into main EP concepts while also providing the opportunity to improve scientific writing skills. Excerpts of the EP syllabus illustrate how the course teaches students to both acquire disciplinary expertise and write as experts in their field: "Students are expected to read carefully and for meaning and be able to demonstrate their mastery of the assigned texts through well-written cogent quizzes, examinations, and lab reports." Further, upon completion of EP, students should be able to "Compile, interpret, and integrate scholarly material related to physiological responses to exercise" and "Discuss through written and oral work how the healthy human body uses intricate communication to coordinate physiological functions during exercise."

During Circles, students participate simultaneously in collaborative discussion and written critique, applying a method we developed and termed "post-outlining" (Kramer, 2016): the writer reads aloud, and then all the peers underline key ideas, bracket topic sentences, and note in the margins of the draft the function of each paragraph, for example, "counterargument" or "evidence for supporting idea," all the while discussing what they are annotating and why, and how the annotated ideas compare with the prompt and the writer's intentions. This post-outlining method is supported indirectly by the findings of Bui and Kong (2019), whose study of written and oral feedback indicates the importance of both. Their data show that students' oral feedback includes more meaning and global issues, while written feedback includes more comments about surface or local issues. Their data also show the need for peers to both receive written critiques and to write down what their peers are saying during discussion, because "written feedback appeared to have a better chance of being incorporated in the later drafts compared to the oral feedback"

(Bui & Kong, 2019, p. 383). This reinforces the Circles model of written annotation coupled with discussion and continual note-taking during discussion.

Circle students are taught principles of effective peer review, as Harris (2014), Kramer (2016), Geithner and Pollastro (2016), and others recommend. Previous research shows that training students on how to do peer review improves both the quantity and the quality of feedback students give to one another. With training, students demonstrate greater facility in commenting on global features of their peers' writing and offering specific feedback for improvements (Zhu, 1995). With effective implementation of peer review, peer response has been shown to include many of the same lexical features as quality response from instructors (Anson & Anson, 2017). Additionally, when peer review is scaffolded with training, well-constructed rubrics, and incentives to take peer reviews seriously, the aggregate of at least four student evaluations of a peers' writing has been shown to be comparable to an instructor's evaluation (Cho, Schunn, & Wilson, 2006). Without this training, as previously noted, we have noticed that peer review can, in some cases, do more harm than good.

Students also learn about and reflect on collaboration, which speaks to Coffey et al.'s (2017) recommendations about the need to "negotiate team communications" (p. 149). While the peers in a Circle typically are not a writing team, in the sense of multiple members working on one group project, they are simultaneously working on the same kind of individual writing projects, and for the same disciplinary course. As peer review teams and disciplinary peers, then, they must learn to navigate the dynamics of team communication. Accordingly, facilitators train students in practical methods of generating helpful feedback, then guide them through asking readerly questions and annotating each other's drafts, the prompt for the assignment, or exemplars (Kramer, 2016).

Circle facilitators represent a diversity of disciplinary backgrounds and experience levels united by a singular method (post-outlining) and a common course calendar. Facilitators issue pass/fail, labor-based grades based on attendance and production. As Inoue (2015) writes, labor-based assessment is more equitable than standards-based systems (p. 84); Inoue points to the removal of concern over grades as a means of encouraging students' willingness to explore and discuss ambiguities, reflect, innovate, and take chances, all of which are markers of critical thinking and essential to productive peer review. This labor-based approach also helps facilitators avoid or climb down from perceived positions of disciplinary authority by their Circle students. Indeed, in a well-functioning Circle, students perceive themselves and their peers as having more disciplinary authority than their facilitator, and the students become comfortable attempting to develop what Beaufort (2007) terms "expert insider prose" (p. 19).

While they strive to avoid a perception of disciplinary authority, facilitators do establish and maintain a practical authority over the structure of discussion in the Circle. Facilitators are hired by CWAC and trained in the post-outlining method so that they in turn can train their Circle students in the post-outlining process as a standard means to effective peer review. Indeed, the facilitator's primary function is to support discussion within the group, specifically, to guide communication among the student group of disciplinary peers through the post-outlining process. The facilitators' role is defined negatively as much as positively, specifically, not only by what they know or do but also by what they do not know or do. While they are familiar with the WID course's assignments and requirements—genre elements, formats, and other features—facilitating instructors do not have the academic training, research experience, or disciplinary knowledge of WID faculty. Thus, while different WID faculty might bring different instructional ideologies, methods, or means to bear within the same WID course, facilitators explicitly bring the same disciplinarily agnostic practice, i.e., post-outlining, to every Circle.

Methods

This comparative, mixed methods study included a quantitative analysis of research proposals written by RMW students without Circles (n=38) in 2011-12 and with Circles (n=39) in 2015; we used these data sets

because we began the research in 2015, and 2011-12 was the last year RMW was taught without WID enhancements. The quantitative analysis also assessed lab reports written by EP students without Circles (n=34) in spring 2016 and lab reports written with Circles (n=69) in fall 2016. All EP students had previously enrolled in a Circle during RMW in their second year. Finally, we conducted a qualitative assessment of the reflections students wrote as part of their Circles with RMW (n=38) in 2015 and with EP (n=52) in fall 2016.

After receiving approval through the Institutional Review Board, we gathered electronic versions of the artifacts, which then were stripped of any information that might indicate the identity of the author or the semester in which they were produced. Each artifact was assigned a number and a pseudonym. A research assistant created and maintained a secure chart of the names, numbers, and pseudonyms. Within the sets of RMW and EP artifacts, the research assistant randomly commingled artifacts so that the coders could not determine the semester in which each was written.

Analysis

To assess the lab reports from EP, we built upon the rubric we had designed for a previous RMW assessment (Kramer et al., 2019), modifying some wording in each category to match the requirements of lab reports in EP as opposed to the research proposals assessed in RMW (Table 1).

Table 1: Assessment Rubric for EP Lab Reports

	4 - Highly developed	**3 - Developed**	**2 - Emerging**	**1 - Initial (attempts)**
Organiza-tion of Ideas & Complexity of Thought within the Discipline	Uses disciplinary conventions of logical & systematic organization, with effective transitions. Methods concisely identify all data & techniques in chronological order. Results contextualize data objectively & effectively. Discussion interprets the importance of results, from the general topic through to conclusionary applications.	Uses disciplinary conventions of organization, with mostly effective transitions. Methods identify all data & techniques in chronological order. Results contextualize data clearly & objectively. Discussion interprets results, from the general topic through to conclusionary applications.	Uses some disciplinary conventions of organization, including some transitions, in a piecemeal or mechanical progression. Methods identify data & techniques. Results contextualize data mostly objectively. Discussion interprets results & includes some conclusions.	Minimal attempt to organize; some content misplaced. Methods incomplete. Some results reported; may be subjective. Discussion attempted.
Format, Tone, & Style	>90% disciplinary, concise language. Paraphrases evidence. Integrates sources elegantly, & all sources cited accurately in APA. Includes all: intro,	About 75% disciplinary, concise language. Privileges paraphrase over quotes, & all sources cited accurately in APA. Includes all: intro, methods, results,	<50% disciplinary, concise language. Privileges paraphrase over quotes of 3 words or less. All sources cited. Includes all: intro, methods, results,	<25% disciplinary language. Privileges quotes over paraphrase. Includes some citations. Includes only some required elements; may include improper ones.

	methods, results, discussion, conclusion, & references page. Results presented purposefully & concisely. Past tense; uses abbreviations appropriately & explained; subjunctive tense; uses "we" correctly. Nearly free of errors.	discussion, conclusion, & references page. Results are easily interpreted by the reader. Past tense; abbreviations explained; no subject names, "you," & "I." Uses "we" correctly. Few error patterns.	discussion, conclusion, & references page. Results are presented distinctly. Mostly past tense; most abbreviations explained; mostly avoids subject names, "you," & "I." Some error patterns impact meaning.	Results presented ineffectively. Sometimes past tense; some abbreviations explained; mostly subject names & 2nd-person pronouns. "We" incorrectly. Many errors that impact meaning.
Intellectual Discovery	Solidly situates the hypothesis & purpose within the discipline, in the context of a pertinent, well-defined gap in knowledge; demonstrates the importance of this study.	Situates the hypothesis & purpose within the discipline, in the context of a gap in knowledge; demonstrates the relevance of this study.	Discusses the hypothesis & purpose, in terms of the discipline generally; offers a broad context of knowledge.	Attempts to form a hypothesis & purpose; offers some perception of the literature; may make unsupported connections.
Synthesis & Analysis of Evidence	Eloquently summarizes, analyzes, & synthesizes evidence, identifying assumptions & conclusions that impact this study in terms of the theoretical framework. Accurately interprets how limitations & sources of error impact study & integrates how these suggest further research.	Skillfully summarizes, analyzes, & synthesizes evidence, identifying some assumptions that impact this study in terms of the theoretical framework. Limitations, sources of error, & suggestions for further research are clearly distinguished.	Summarizes, analyzes, & begins to synthesize evidence to support ideas in the context of this study. Limitations, sources of error, & suggestions for further research are discussed.	Summarizes & attempts to analyze evidence; may not clearly support ideas. Attempts to discuss limitations, sources of error, and/or suggestions for further research.
Theoretical Frame-work	Clear & concise explanations of key terms, concepts, theories, or principles & their implications in the context of the research question.	Explains key terms, concepts, theories, or principles in the context of the research question.	Discusses key terms, concepts, theories, or principles.	Mentions key terms, concepts, theories, or principles; may be misidentified.

Using the relevant rubric, each reader participated in extensive norming sessions by coding and discussing several artifacts not included in the study. Based on discussion during norming, we revised the wording of the RMW and the EP rubrics. The overall inter-rater reliability (IRR), specifically, raters agreeing within 0 to 1 point, was 95% for the norming sessions. Following norming, seven readers coded the 77 blinded, commingled RMW research proposals and the 103 blinded, commingled lab reports from EP. Each artifact was coded by two readers, each of whom worked with a clean copy to independently assign scores on a 4-point scale for each metric on the rubric, plus a holistic score. A third reader was added when holistic scores differed by one point or more. For the RMW artifact coding, 19% triggered a third reader; whereas fewer than 3% of the EP artifacts triggered a third reader. Scores from each set of readers for a given artifact were averaged. We then compiled the results and analyzed them.

In the analysis of rubric scores, although we coded two lab reports from spring 2016 and four from fall 2016, we included only the third EP report, on which students demonstrated their peak writing performance (n=24 for spring 2016 and n=17 for fall 2016), in order to compare the results authentically with data from the single RMW research proposal. Significance differences in mean scores between the various semesters were analyzed using an unpaired *t*-test comparing different students' scores. Statistical significance was set at $\alpha = .05$, and values are represented as means ± standard error, unless otherwise noted. Effect sizes were analyzed by Hedge's *g* formula, using mean values and pooled weighted standard deviations. Values greater than 0.75 were indicative of a large effect size.

The written reflections from both the RMW and EP Circles were also analyzed. Two readers separately coded each reflection artifact. Thematic analysis was used to create some preliminary codes, using highlighters to separate data into groups for further analysis and description (Glesne, 2011). Descriptive, low inference coding was used during the first reading of the reflections. During the second reading, the first and second readings of the data were synthesized to find emerging themes and descriptions that aligned with the study research questions. The content of each theme was analyzed using constant comparison, resulting in the final interpretations of the data gathered.

Results

Students who participated in weekly Circles in the RMW course scored significantly higher ($p < .05$) in all five rubric metrics compared to students who took RMW prior to the implementation of Circles (Figure 1). These improvements were particularly noticeable for the course writing learning outcomes of *Organization of Ideas & Complexity of Thought* (+40%) and *Format, Tone, & Style* (+48%) (Table 2).

The scores from the two RMW semesters were compared to those from the two EP semesters. Compared to the RMW research proposals, scores were generally higher for the EP lab reports, particularly for the critical thinking learning outcomes of *Intellectual Discovery* and *Theoretical Framework*. Specifically, from the RMW without Circles to the EP without Circles, improvements were observed in all five rubic metrics: *Organization of Ideas* (+39%), *Format, Tone, & Style* (+58%), *Intellectual Discovery* (+48%), *Synthesis & Analysis* (+25%), and *Theoretical Framework* (+57%). However, improvements in the EP with Circles compared to RMW with Circles were observed primarily in the critical thinking learning outcomes *Intellectual Discovery* (+45%) and *Theoretical Framework* (+29%); whereas the other three categories, related to writing, remained relatively unchanged (Figure 1). Circles were particularly effective for improving students' *Intellectual Discovery* in the EP with Circles, which increased by 23% compared to the EP without Circles (Table 2).

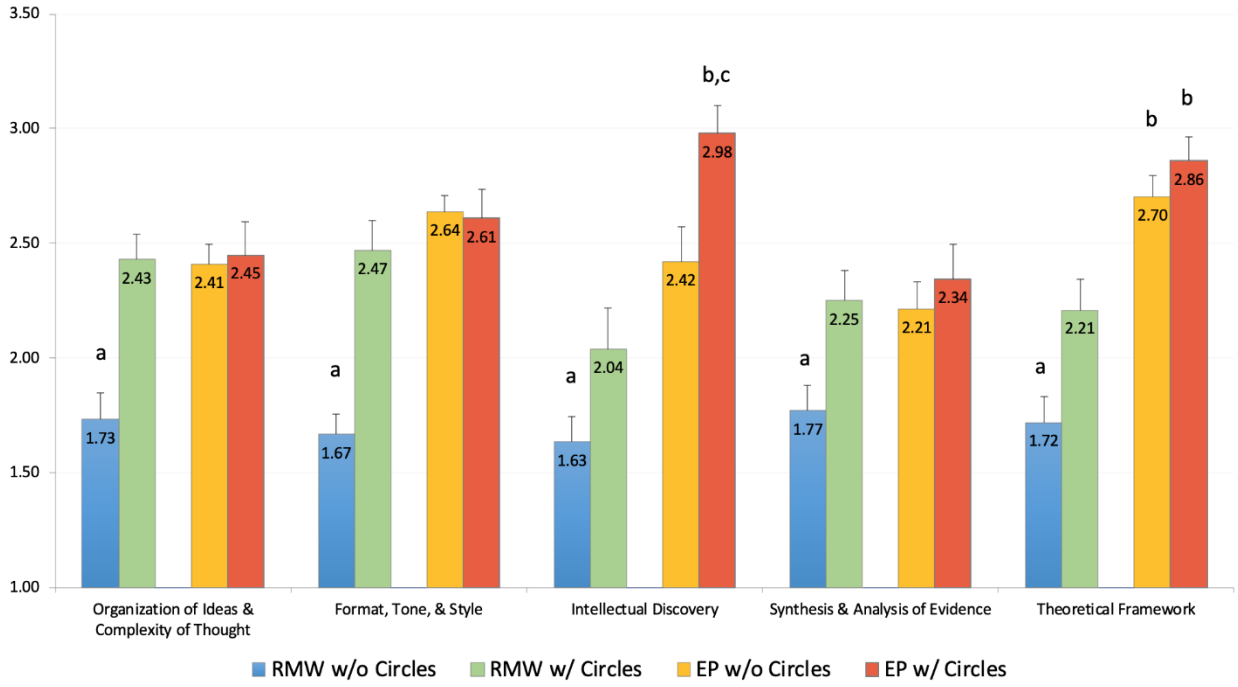

Figure 1: Mean Scores by Rubric Metrics, Chronologically for Research Methods & Writing and Exercise Physiology Courses Without and With Circles.

All scores are based on a 4-point scale.
a = significantly different from all other semesters ($p < .05$)
b = significantly different from RMW w/ Circles ($p < .05$)
c = significantly different from EP w/o Circles ($p < .05$)

Table 2: Effect Sizes (and Deltas) Comparing Four Semesters of Mean Scores

	RMW w/ Circles	EP w/o Circles	EP w/ Circles
Organization of Ideas			
RMW w/o Circles	**0.99 (+40%)**	**1.08 (+39%)**	**1.05 (+41%)**
RMW w/ Circles	-	0.03 (-1%)	0.03 (+1%)
EP w/o Circles	-	-	0.08 (+2%)
Format, Tone, & Style			
RMW w/o Circles	**1.16 (+48%)**	**2.09 (+58%)**	**1.76 (+56%)**

RMW w/ Circles	-	0.24 (+7%)	0.18 (+5%)
EP w/o Circles	-	-	0.07 (-1%)
Intellectual Discovery			
RMW w/o Circles	0.46 (+26%)	**1.11 (+48%)**	**2.13 (+82%)**
RMW w/ Circles	-	0.36 (+18%)	**0.95 (+45%)**
EP w/o Circles	-	-	**0.83 (+23%)**
Synthesis & Analysis			
RMW w/o Circles	0.66 (+28%)	0.68 (+25%)	**0.85 (+32%)**
RMW w/ Circles	-	0.08 (-3%)	0.09 (+3%)
EP w/o Circles	-	-	0.21 (+6%)
Theoretical Framework			
RMW w/o Circles	0.65 (+29%)	**1.59 (+57%)**	**1.79 (+66%)**
RMW w/ Circles	-	0.67 (+22%)	**0.85 (+29%)**
EP w/o Circles	-	-	0.34 (+6%)

Large effect sizes of >0.75 are highlighted in bold.

Through our qualitative analysis of the reflections that students wrote at the conclusion of their Circles, these themes emerged as most important, as measured by their frequency: collaboration, writing process, transfer of knowledge, and kinesiology as genre. The reflections of both RMW and EP students include evidence of knowledge building in each of the four themes. However, there are notable distinctions between the reflections of the second-year RMW students and the reflections of the fourth-year EP students, primarily around issues of peer authority, and these are the issues on which we focus our discussion.

Discussion

Peer Review in STEM Courses Over Time

Previous research has demonstrated some benefits of peer review in teaching students how to write in STEM courses, during which teaching content often takes precedence over the development of writing skills (Brieger & Bromley, 2014; Geithner & Pollastro, 2016; Kramer et al., 2019). A central purpose of this study,

therefore, was to examine how the impact of peer review is realized over time within a STEM major. We hypothesized that the benefits of peer review manifest in different ways at different times in students' academic and psycho-socio-epistemic development. Through quantitative coding of student writing artifacts according to a five-metric rubric, accompanied by qualitative analysis of the students' reflections, we found that the implementation of Circles as a companion to a WID course impacts student learning differently at the start versus the end of their disciplinary studies.

There are four key findings of this study. First, Circles are effective at improving all areas of writing and critical thinking in the lower-division course, and their impact is most dramatic on the learning outcomes closely tied to writing and the writing process. Second, Circles in the upper-division course are most effective at improving learning outcomes associated with critical thinking. Third, when comparing scores between the lower-division and upper-division courses, we see that writing and writing process scores were relatively unchanged, but the scores for critical thinking improved significantly. Fourth, when comparing their reflections, we see that only the upper-division students both gain and grant their peers disciplinary authority as peer reviewers. We are encouraged to see writers improving throughout their college career. Students are learning in their lower-division WID course the mechanics of writing as kinesiologists, while also improving as critical thinkers. By the time students take the final or near-final course in their major, their growth in mechanical writing skills may have begun to plateau but their growth as critical thinkers and disciplinary analysts continues to improve.

For a course that requires students to produce scientific writing, our study shows that a companion Circles course confers widespread benefits. Toward the beginning of students' progress through the discipline, there is significant improvement in all the learning outcomes. During the lower-division RMW course, which is designed to train students to write in the major, the largest improvements were in written communication, specifically, the learning outcomes of *Organization of Ideas* and *Format, Tone, & Style*.

This study found distinct differences in the impact of Circles in the lower-division RMW vs. the upper-division EP course, which students typically take in their fourth year. The EP students had all previously taken RMW with Circles and were already familiar with the collaborative practice of peer review via post-outlining. The rubric metric that demonstrated statistically significant improvement during EP was *Intellectual Discovery*, a core learning outcome in the development of disciplinary critical thinking skills. Our study shows that in the upper-division course, Circles were less important in honing general writing skills, but rather augmented the students' ability to articulate specific bodies of knowledge within the context of the discipline. We suggest that this illustrates that RMW and the connected Circles at that stage are achieving their learning outcomes for preparing students to write in the major. EP comes at or near the end of students' college careers; therefore, we are encouraged that the students are growing primarily as critical thinkers in their discipline at that point.

The Role of Authority in Peer Review

Our data show that the benefits of structured, iterated, facilitated peer review manifest in different ways at different times as students progress through their major. We argue that the primary cause of these differences is the extent to which students are able to perceive both themselves and each other as authentic, authoritative sources of knowledge in their discipline. Looking at the student writing for both the upper- and lower-division kinesiology courses as well as the same students' reflections during their Writing Circles, we see that peer feedback which students give and receive in Circles impacts their writing differently depending on how they authorize themselves and their peers as sources of knowledge.

Examining connections among audience awareness, authenticity, and authority in peer review groups, Keating (2019) argues that group efficacy depends on "the extent to which students felt that their peers were authorized, by each other and their instructors, to give feedback" (p. 57). Our findings suggest that students at an early stage of disciplinary studies feel more authorized to give, receive, and act on feedback related to

organization and format, tone, and style than on feedback related to intellectual discovery and theoretical framework. Why? Consider that the post-outlining—annotating and collaborative discussing—that occurs in Circles is an effective means to decentralize the authority of the instructor and redistribute it to student Circle peers. This follows Bruffee's (1984) observation that during peer review, the author of the prompt (the instructor), is not present to explain the prompt or confirm or deny interpretations, which makes everyone in the Circle an equally potential source of insight.

Only the fourth-year students' reflections revealed evidence of learning to trust each other as disciplinary peers. The reflections by both sets of students included similar evidence of trust and community building more generally, as writers and peer students, which illustrates that at all stages of a student's college career, peer review can be an effective way to build trust and community among majors. Van Gennip et al. (2010) demonstrate that participating in a peer assessment intervention has a positive effect on a number of interpersonal variables. These include an increase in feelings of psychological safety (i.e., okay to take interpersonal risks in a group of people) and lower value diversity (i.e., less disagreement about a team's task, goal, or mission); furthermore, students express more trust in their peers as assessors of writing. We saw references to similar variables throughout our students' reflections. As RMW student Fry reflects, "I felt that none of us judged each other, so it was never intimidating when we had to share our work." And RMW peer Jon writes that "I found my Writing Circle to be a very calm, open, relaxed place where everyone was there to help each other, not judge each other."

While second-year students may not understand themselves or their peers as reliable sources of theoretical or lexical conventions of their discourse communities, they do come to regard one another as dependable arbiters of the organizational and stylistic conventions detailed in scaffolded prompts, even if only that. The many reflection comments of this nature, such as RMW student Fry appreciating his Circle peers' help with "what the guidelines were for our assignments," read as though they could have been written by students of any discipline, that is, there is no indication that the second-year students sought out or appreciated their peers' insider knowledge about kinesiology, but rather, were seeking feedback on the norms of writing.

Conversely, the fourth-year students frequently reflect on their peers' contributions in ways that indicate they view them as authorities in the discipline who are helping each other develop what Beaufort (2007) describes as "expert insider prose" (p. 19). As EP student Mari writes: "having people with the background knowledge gave validity to their suggestions." We argue that this reliance on peers for disciplinary knowledge correlates with the critical thinking gains experienced by these same fourth-year students. Because they grant each other more kinesiology-specific authority at this stage of their college careers, they are more likely to have disciplinary conversations in Circles and therefore more likely to grow as thinkers in their discipline.

In the sciences, such shared inquiry in a writing setting effectively doubles as a study group for the content of lab reports. We see evidence of this in our data, as students express gratitude for their peers' help. EP student Milo reflects: "Starting out, my methods section was all over the place, however, towards the last two lab reports, I was able to condense and clarify what exactly we did and if someone were to recreate our lab [they] could do it based on my methods," and, "There were many times I…forgot to mention an important definition…Having my peer point that out helped me include everything I need to." We concur with Geithner and Pollastro's (2016) assertion that peer review "provided students with opportunities to work effectively and collaboratively with each other and to critically discuss and contemplate new ideas...to experience the sense of being part of a scientific learning community" (p. 43). Through peer review, students are encouraged to rely on each other as valuable resources in their specific disciplines, introducing the kind of collaboration expected in a professional setting.

Limitations

We note three potential limitations which point to continued research. First, this study did not use randomly assigned participants but rather four intact groups of participants; however, the two RMW groups were similar in that they both included mostly students in their second year, and the two EP groups included mostly students in their fourth year. Second, the lower-division RMW students had yet to specialize in one of the three tracks in kinesiology offered at our college, suggesting that our results are more generalizable than they might appear to be: while the research proposals were all examples of STEM writing—scientific research proposals with empirical research questions and testable hypotheses—the writers were not all students with futures in STEM; rather, they represented a cross-section of the undergraduate population and the wide variety of rhetorical experiences typical of students early in their college career. By contrast, students in the upper-division EP course had all chosen the STEM-oriented Health & Human Performance track of the kinesiology major. Third, our study is comparative rather than longitudinal because we did not follow the same individuals from second to fourth year. Our next steps could be a study of the same design involving the other two tracks of upper-division kinesiology majors, or a longitudinal study. For this current study, the data are comparative because all the participants were kinesiology majors; additionally, all the EP participants had taken RMW during their second year, and all with a connected Circle in both RMW and EP.

Conclusions

A social constructionist perspective sees collaborative learning as "helping students converse with increasing facility in the language of the communities they want to join" (Bruffee, 1993, p. 73). As vehicles for collaborative learning, disciplinary peer review groups can decentralize authoritative barriers to entry that frustrate the efforts of WID students seeking to join discourse communities. Just how peer review manifests, however, depends on when and how it appears in students' academic and psycho-social development.

Toth & Aull (2014) found that the progression of writing skills throughout college is uneven, as mechanics improve earlier than critical analysis skills: the latter undergo their most significant improvement in the later stages of the major, as mechanics level off; conversely, the most significant improvement in mechanics occurs in the early stages, as critical-analysis skills languish comparatively unchanged. In our study, however, both writing and critical thinking learning outcomes improved in the second-year course, which we attribute to the WID course enhancements and the connected weekly peer review in Circles. We argue that this points to the effectiveness of placing WID supports early in students' careers: in that way, students can learn how to write in their discipline and carry that learning throughout their studies, rather than waiting until a writing-intensive fourth-year course, as is often the case in STEM majors.

Based on our results, we propose that how the peer review is structured and presented matters. Intentionally teaching students to view each other as disciplinary authorities while they peer review, at all stages of their college career, will help them improve more as critical thinkers. Our study adds support for Keating's (2019) musing that "making the case to students about peer review as a worthwhile tool will mean continuing to research the enactment of collaborative kinds of authority in classroom discourse" (p. 76). Data from our fourth-year students offer evidence for this enactment, as the peers reflect on trusting each other to offer expert insider guidance, which elevates their collaborative discussion. Only the upper-division students reflected on each other's authority within the discipline, which we view as connected to the students' improvements in critical thinking at that stage. When the peers are trusted as disciplinary authorities, everyone at the table thinks and learns more deeply.

References

Anson, Ian G., & Anson, Chris M. (2017). Assessing peer and instructor response to writing: A corpus analysis from an expert survey. *Assessing Writing*, *33*, 12-24. https://doi.org/10.1016/j.asw.2017.03.001

Armstrong, Sonya L., & Paulson, Eric J. (2008). Whither "peer review"? Terminology matters for the writing classroom. *Teaching English in the Two-Year College*, *35*(4), 398-407. https://eric.ed.gov/?id=EJ792050

Baker, Kimberly M. (2016). Peer review as a strategy for improving students' writing process. *Active Learning in Higher Education*, *17*(3), 179-192. https://doi.org/10.1177/1469787416654794

Beaufort, Anne. (2007). *College writing and beyond: A new framework for university writing instruction.* Utah State University Press.

Brieger, Katharine, & Bromley, Pam. (2014). A model for facilitating peer review in the STEM disciplines: A case study of peer review workshops supporting student writing in introductory biology courses. *Double Helix*, *2*, 1-10. https://wac.colostate.edu/docs/double-helix/v2/brieger.pdf

Bruffee, Kenneth A. (1984). Collaborative learning and the 'conversation of mankind'. *College English, 46*(7), 635-652. https://www.jstor.org/stable/376924

Bruffee, Kenneth A. (1993). *Collaborative learning: Higher education, interdependence, and the authority of knowledge*. Baltimore: Johns Hopkins University Press.

Bui, Gavin, & Kong, Amy. (2019). Metacognitive instruction for peer review interaction in L2 writing. *Journal of Writing Research, 11*(2), 357-392. https://doi.org/10.17239/jowr-2019.11.02.05

Cho, Kwangsu, & Schunn, Christian D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, *48*(3), 409-426. https://doi.org/10.1016/j.compedu.2005.02.004

Cho, Kwangsu, Schunn, Christian D., & Wilson, Roy W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, *98*(4), 891-901. https://doi.org/10.1037/0022-0663.98.4.891

Coffey, Kathleen M., Gelms, Bridget, Johnson, Cynthia C., & McKee, Heidi A. (2017). Consulting with collaborative writing teams. *The Writing Center Journal*, *36*(1), 147-182. http://www.jstor.org/stable/44252641

Deans, Thomas. (2017). One-credit writing-intensive courses in the disciplines: Results from a study of four departments. *Across the Disciplines*, *14*(1), 1-25. https://wac.colostate.edu/atd/articles/deans2017.cfm

DiPardo, Anne, & Warshauer, Freedman. (1988). Peer response groups in the writing classroom: Theoretic foundations and new directions. *Review of Educational Research*, *58*(2), 119-149.

Finkenstaedt-Quinn, Solaire A., Snyder-White, Emma P., Connor, Megan C., Gere, Anne Ruggles, & Shultz, Ginger V. (2019). Characterizing peer review comments and revision from a writing-to-learn assignment focused on Lewis structures. *Journal of Chemical Education*, *96*(2), 227-237. https://doi.org/10.1021/acs.jchemed.8b00711

Finkenstaedt-Quinn, Solaire A., Polakowski, Noelle, Gunderson, Brenda, Shultz, Ginger V., Gere, Anne Ruggles (2021). Utilizing peer review and revision in STEM to support the development of conceptual knowledge through writing. *Written Communication*. https://doi.org/10.1177/07410883211006038

Geithner, Christina A., & Pollastro, Alexandria N. (2016). Doing peer review and receiving feedback: Impact on scientific literacy and writing skills. *Advances in Physiology Education, 40*(1), 38-46. https://doi.org/10.1152/advan.00071.2015

Gerdeman, R. Dean, Russell, Arlene A., & Worden, Kelly J. (2007). Web-based student writing and reviewing in a large biology lecture course. *Journal of College Science Teaching*, *36*(5), 46-52.

Gere, Anne Ruggles. (1987). *Writing groups: History, theory, and implications*. Southern Illinois University Press.

Gere, Anne Ruggles. (2019). Conclusion. In Anne Ruggles Gere (Ed.), *Developing writers in higher education: A longitudinal study* (pp. 313-328). University of Michigan Press. https://www.jstor.org/stable/j.ctvdjrpt3

Glesne, Corrine. (2010). *Becoming qualitative researchers: An introduction* (4th ed.). Pearson.

Graham, Joan. (1992). Writing components, writing adjuncts, writing links. In Susan H. McLeod, & Margot Soven. (Eds.), *Writing across the curriculum: A guide to developing programs* (pp. 78-93). Sage Publications.

Harris, Muriel. (2014). Afterword: A non-code: Including writing centered student perspectives for peer review. In Steven J. Corbett, Teagen E. Decker, & Michelle LaFrance, (Eds.), *Peer pressure, peer power: Theory and practice in peer review and response for the writing classroom* (pp. 277-288). Fountainhead Press.

Inoue, Asao B. (2015). *Antiracist writing assessment ecologies: Teaching and assessing writing for a socially just future.* Colorado: The WAC Clearinghouse; South Carolina: Parlor Press. https://wac.colostate.edu/books/perspectives/inoue/

Keating, Benjamin. (2019). "A good development thing": A longitudinal analysis of peer review and authority in undergraduate writing. In Gere, Anne R., & Silver, Naomi (Eds.), *Developing writers in higher education: A longitudinal study* (pp. 56-79). University of Michigan Press. https://www.jstor.org/stable/j.ctvdjrpt3

Kramer, Tereza. (2016). Writing circles: Combining peer review, commitment, and gentle guidance. *WLN: A Journal of Writing Center Scholarship, 40*(7-8), 20-27. https://wlnjournal.org/archives/v40/40.7-8.pdf

Kramer, Tereza Joy, Zeccardi, Joe, Concepcion, Rebecca, Emhoff, Chi-An, Miller, Steve, & Posell, Krista V. (2019, December 28). WID course enhancements in STEM: The impact of adding "writing circles" and writing process pedagogy. *Across the Disciplines, 16*(4), 26-37. http://wac.colostate.edu/docs/atd/articles/krameretal2019.pdf

Pelaez, Nancy J. (2002). Problem-based writing with peer review improves academic performance in physiology. *Advances in Physiology Education*, *26*(3), 174-184. https://doi.org/10.1152/advan.00041.2001

Reddy, Krishneel, Harland, Tony, Wass, Rob, & Wald, Nave. (2021). Student peer review as a process of knowledge creation through dialogue. *Higher Education Research & Development*, *40*(4), 825-837. https://doi.org/10.1080/07294360.2020.1781797

Reynolds, Julie A., Thaiss, Christopher, Katkin, Wendy, & Thompson, Robert J. (2012). Writing-to-learn in undergraduate science education: A community-based, conceptually driven approach. *CBE - Life Sciences Education*, *11*(1), 17-25. https://doi.org/10.1187/cbe.11-08-0064

Rivard, Lé Onard P. (1994). A review of writing to learn in science: Implications for practice and research. *Journal of Research in Science Teaching*, *31*(9), 969-983. https://doi.org/10.1002/tea.3660310910

Toth, Christie, & Aull, Laura. (2014). Directed self-placement questionnaire design: Practices, problems, possibilities. *Assessing Writing*, *20*, 1-18. https://doi.org/10.1016/j.asw.2013.11.006

van Gennip, Nanine A. E., Segers, Mien S. R., & Tillema, Harm H. (2010). Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction, 20*(4), 280-290. https://doi.org/10.1016/j.learninstruc.2009.08.010

Walvoord, Mark E., Hoefnagels, Mariëlle H., Gaffin, Douglas D., Chumchal, Matthew M., & Long, David A. (2008). An analysis of calibrated peer review (CPR) in a science lecture classroom. *Journal of College Science Teaching*, *37*(4), 66-73.

Yalch, Matthew M., Vitale, Erika M., Ford, J. K. (2019). Benefits of peer review on students' writing. *Psychology Learning & Teaching*, *18*(3), 317-325. https://doi.org/10.1177%2F1475725719835070

Zhu, Wei. (1995). Effects of training for peer response on students' comments and interaction. *Written Communication*, *12*(4), 492-528. https://doi.org/10.1177/0741088395012004004

Contact Information

Tereza Joy Kramer
Assistant Professor
California Northstate University
Email: tereza.joy.kramer@cnsu.edu

Joe Zeccardi
Adjunct Associate Professor
Saint Mary's College of California
Email: jaz4@saintmarys-ca.edu

Chi-An W. Emhoff
Associate Professor
Saint Mary's College of California
Email: cwe1@saintmarys-ca.edu

Claire Williams
Associate Professor
Saint Mary's College of California
Email: cmw9@saintmarys-ca.edu

Robin J. Dunn
Associate Professor
Saint Mary's College of California
Email: rd19@saintmarys-ca.edu

Joshua Rose
Instruction and Reference Librarian
College of Alameda
Email: jrose@peralta.edu

Complete APA Citation

Kramer, Tereza Joy, Zeccardi, Joe, Emhoff, Chi-An W., Williams, Claire, Dunn, Robin J., & Rose, Joshua. (2022, February 18). How timing and authority in peer review impact STEM students: A comparative assessment of writing and critical thinking in kinesiology courses. *Across the Disciplines*, 18(3/4), 305-319. https://doi.org/10.37514/ATD-J.2022.18.3-4.06