3-15-2009

Faculty Should Consider Peer Review as a Means of Improving Students' Scientific Reasoning Skill

Briana Eileen Timmerman
University of South Carolina - Columbia, timmerman@sc.edu

Denise Strickland
University of Virginia - Main Campus

Publication Info

Journal of the South Carolina Academy of Science, ed. Hans-Conrad zur Loye, Volume 7, Issue 1, 2009, pages 1-7.

Faculty Should Consider Peer Review as a Means of Improving Students' Scientific Reasoning Skills

Briana Timmerman[a], Denise Strickland[b]

[a]SC Honors College and Department of Biological Sciences, University of South Carolina, Columbia SC 29208; timmerman@schc.sc.edu
[b]Department of Educational Studies, University of South Carolina, Columbia SC 29208

Received March 15, 2009

The following article provides highlights of a larger body of work on the impact of peer review on undergraduate biology students' scientific reasoning, scientific writing and attitudes. Results suggest that undergraduates, even freshman can be effective peer reviewers and that peer review improves scientific writing, content knowledge, and scientific reasoning skills. Students report peer review to be a beneficial experience both within the course and in terms of helping them to develop as aspiring scientists. Science faculty are therefore encouraged to consider incorporating peer review as a regular part of instruction.

Introduction and Rationale

Despite large volumes of literature on the benefits of reformed curricula for improving student learning [1, 2], pedagogical revolution has been slow to occur in many higher education institutions. We suggest that this is due in part to the large time and/or resource investments required for faculty to adapt and incorporate innovative strategies. Many pedagogical innovations require unfamiliar technologies or methodologies e.g. [3], and the common lack of pedagogical support for higher education faculty leaves many instructors to simply teach as they were taught [4, 5]. Peer review is considered a productive learning experience for graduate students [6, 7], and we suggest that it is likely valuable and effective for improving undergraduates' scientific reasoning abilities, scientific writing, and attitudes regarding science as well.

There are several major motivations for including peer review in science classrooms. First, the ability to critique and evaluate the quality of scientific claims is an important scientific skill in and of itself, but, like other reasoning skills, it must be an explicit component of instruction and students must be given opportunities to practice and improve it as a skill. Second, understanding the role of peer review as a major accountability mechanism and source of credibility and integrity of science knowledge is critical to public confidence in science. Use of peer review in the classroom causes significant gains in undergraduates' knowledge of this critical function of peer review in the scientific community [8]. Third, using peer review allows instructors to incorporate more writing assignments without correspondingly increasing their grading load. Fourth, peer review improves content knowledge [9], writing [10, 11], and, we hypothesize, scientific reasoning skills.

Peer review is not a new pedagogical technique for some science faculty and multiple online tools exist. For example, the *Calibrated Peer Review*™ system (http://cpr.molsci.ucla.edu) has been incorporated into over 4000 courses (including the authors') across a wide range of science disciplines[12]. Other online peer review systems include Scaffolded Writing and Reviewing in the Discipline (SWORD) (http://www.lrdc.pitt.edu/schunn/sword/index.html) and Waypoint Outcomes which runs through Blackboard (www.subjectivemetrics.com). All these systems allow faculty to engage their students in peer review anonymously and outside of class time.

Regardless of the method of implementation, the central impact of peer review on the student can fall into one of two categories: formative feedback as a mechanism for learning, or summative grading. While both provide the benefits of engaging students in the evaluation of scientific thoughts and writing, we feel that formative feedback stimulates greater learning because students can apply new ideas gained from the experience directly and immediately. Contrary to many instructors' and students' initial concerns, previous research has shown that undergraduate peers can be valid and reliable reviewers [13], regardless of academic strength [14].

Student perceptions of Peer Review

The *Calibrated Peer Review*™ (CPR) system was used in Introductory Biology courses for majors at our institution and we measured its impact on students' writing, scientific reasoning and attitudes. Student attitudes were assessed using an anonymous online survey that was administered in several different biology classes over multiple semesters (total n = 1026 students). From the survey data, we know that an overwhelming majority of the students viewed their peer review

experience positively (Figure 1). Among other benefits, students reported that peer review improved their content knowledge, their scientific writing skills (generally as well as directly impacting the assignment at hand), and their critical thinking skills. Interestingly, a vast majority (85%, n = 557) agreed that even just the act of giving feedback was helpful to improving their own work because it stimulated self-assessment (see also [15]. Lastly, students reported that the act of engaging in peer review provided a window into realistic scientific practice and contributed to their development as practicing scientists.

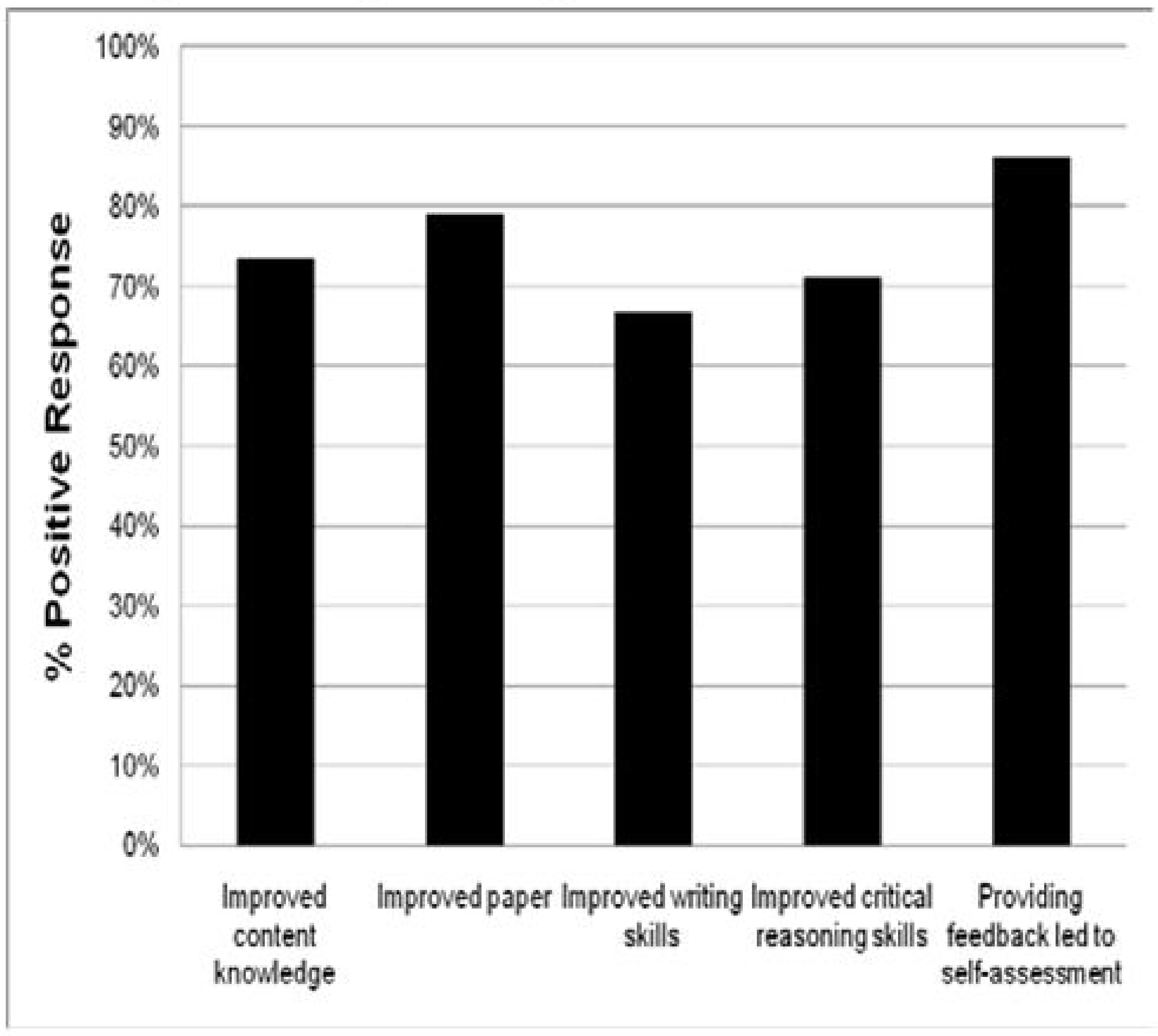

Figure 1. Student perceptions of the role and impact of peer review from surveys of three Introductory Biology courses for majors (total n = 1026 students, number of responses per item ranges from 440 to 998).

While many instructors are sensitive to student opinion, most are equally concerned with how much students are actually learning. Therefore, the effectiveness of the peer review experience was also measured directly using two different mechanisms of measurement: student papers and an independent objective (multiple choice) measure of scientific reasoning skill.

Effect of Peer Review on Student Scientific Writing

When we looked directly at the laboratory reports involved in the peer review process, we primarily wanted to know if peer review actually improved the papers. We found that peers were capable of providing multiple useful feedback items and when students made changes in their papers based on this peer feedback, scores on lab reports increased (Figure 2).

In our experience, two major factors contribute to the usefulness of formative peer feedback: explicit instruction about what constitutes

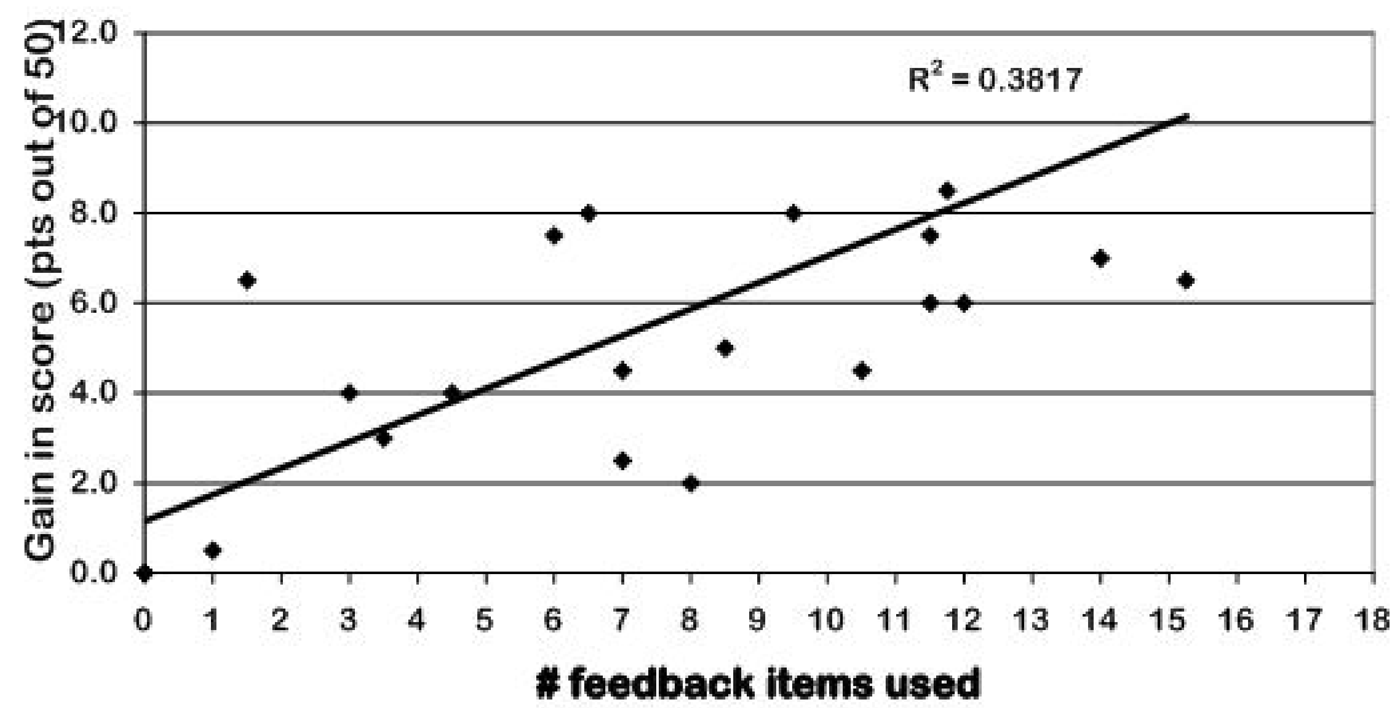

Figure 2. Effect of peer feedback on the quality of Introductory Biology lab reports. Gain is change in total score from draft to final version (n = 22 unique papers).

useful feedback, and student accountability to provide such feedback. A discussion of useful feedback was included in an explanation of the role of peer review in the scientific community and helped students focus on providing constructive criticism rather than making value judgements. A representative handout with feedback examples helped students to understand these criteria (Appendix 1), and help them to avoid overly positive or generic comments, neither of which is useful [16, 17]. Accountability can be ensured by including the review process in the grade for an assignment. One method we used was to randomly select one review per student to be quickly skimmed for useful feedback, resulting in a grade equivalent to a quiz. Therefore, with proper planning, freshman undergraduates have effectively engaged in peer review and used the resulting feedback to improve their scientific writing.

Next we were concerned with how peer review impacted students compared to those who did not engage in peer review. In order to compare performance across multiple classes, we developed the Universal Rubric for Science Writing. We subjected the Rubric to formal reliability testing using three science graduate student raters per paper across three different biology courses (total of 9 raters) and found it to be highly reliable (*generalizability coefficient* = 0.85 (Timmerman *et al* manuscript submitted). When we compared lab reports from introductory biology courses (BIOL 101 and 102) that incorporated peer review with those from a course that did not use peer review (BIOL 301), students in the introductory biology course outperformed students in the upper level course on many of the scientific reasoning criteria (total n = 142 papers, ANOVA $p = 0.001$). In particular, freshman outperformed upperclassman in the areas of data selection, data presentation and use of primary literature despite the fact that students in the upper-level course had more

academic experience and a higher average institutional GPA.

Impact of Peer Review on Scientific Reasoning

For our second measure of the impact of peer review process on scientific reasoning skills we intentionally selected an assessment tool that used mostly physical science scenarios to ensure that results would not be influenced by students; prior biology course enrolment. Biology majors in five different biology courses (total n = 581 students), ranging from freshmen level to seniors, took the *Test of Scientific Reasoning* [18, 19].

As expected, students' scores improved significantly as they gained academic experience (as measured by number of credit hours) with freshman scores averaging 4.89 (on a 12 point scale), seniors 6.07 and sophomores and juniors falling in between (ANOVA $p = 0.011$). When scores were categorized based on the number of peer review experiences in which a student had engaged however, freshman who had engaged in 2 peer review experiences scored significantly higher (average score = 6.82, n = 61) than students who did not engage in any peer review experiences (ave. = 5.22, n = 260) with students who engaged in a single peer review experience scoring intermediately (ANOVA $p = 0.000$).

Conclusions and Recommendations

Based on these results, we recommend that science faculty incorporate peer review into their courses. Peer review has been previously demonstrated to benefit graduate students [8] and these results suggest that it benefits even freshman in large introductory courses. Peer review allows faculty to incorporate or increase student writing in their course (with all its associated benefits) without incurring significant costs in terms of time and effort. It improves scientific reasoning and writing because it provides three to four times the practice at writing and evaluating as well as increases formative feedback while decreasing instructor load. Additionally, students report it to be a worthwhile and productive experience and perceive it to be an important component of their development as future scientists. Based on our experiences of using peer review effectively, we provide the following recommendations to help other faculty incorporate peer review (Table 1).

Table 1. Summary of Recommendations for Implementing Peer Review.

- **Be explicit** in discussing with students the role of peer review in the scientific community as well as its benefits in the classroom.
- **Share research results** with students demonstrating that peers are effective reviewers and that peers can provide useful feedback that improves paper quality if incorporated (e.g. Figure 2).
- Design assignments to **encourage students to provide high quality written feedback**.
    - Explicitly define and discuss the characteristics of useful feedback (Appendix 1) and
    - Use accountability measures which reward students who make honest efforts at providing useful feedback.
- Design assignments to **align assignment criteria, peer review criteria and instructional goals**. Ideally, instructional goals span multiple courses and expectations for student performance are consistently aligned and developed throughout those educational experiences.
- **Make your expectations explicit** and explain the criteria for a writing assignment to students (e.g. Appendix 2). Better yet, use a rubric as a means of defining assignment criteria to students. Descriptions of what constitutes different levels of performance deepens student understanding of the intent of criteria and helps them to provide better feedback to peers as well as better understand the learning goals.

Acknowledgements

This work was made possible by National Science Foundation Award 0410992. Please note that while the authors take responsibility for the integrity of the data and conclusions made in this paper, the peer review program at USC itself is a product of the combined intellects of Susan M. Carstensen, Dr. Laurel Hester, Dr. Michelle Vieyra, and Dr. Kirk A. Stowe. Many thanks to K. Dahlke, and L. Powell for their critical and central roles in the collection and analysis of the survey data. The authors sincerely thank the faculty and especially the graduate teaching assistants of Department of Biological Sciences for their efforts and cooperation over the multiple years required to integrate and assess the impact of peer review in the curriculum.

Appendix 1: Student Handout
How to provide useful feedback

Useful feedback is:

- **specific and concrete,**
- **focuses on the quality** of the author's argument (e.g. are conclusions logical and supported by the evidence/data?) **rather than on the mechanics** of writing, (e.g. spelling or grammar),
- **identifies hidden or implicit assumptions or consequences** of author's ideas.
- tells the writer **why** you think they did or did not meet the criteria**.**
- In sum, useful feedback is likely to result in meaningful revisions or new content being added to the paper.

You will be prompted by the peer review website to provide feedback for three papers written by your peers. The criteria are the same as those in the assignment handout. Your TA will randomly select one of your reviews and grade the quality of the feedback you provide. Useful items earn 1 pt and partially useful items earn 0.5 pts. Non-useful feedback earn 0 points. A review may earn up to 10 points. You may write as many feedback items as you would like, but you must provide at least one piece of feedback in response to each prompt. Reviewing other students' papers may also give you insight into strengths and weaknesses in your own paper. The benefits you will receive from this exercise are directly correlated with the effort you put into it. Below are some examples of what constitutes useful feedback.

Examples:

Feedback item	Useful?	How to improve the feedback
1. Your paper is GREAT! How did you come up with your idea?	NO	Provides no information to the writer on how to improve the paper.
2. At the end of paragraph 2, you say you think this was a sex-linked cross. Is this your hypothesis? What traits do you think the parents had? Why do you think this is the best explanation?	Yes	Feedback has detail on where and why the reviewer was lost. If the writer answers the reviewer's questions, the paper will have a clearer statement of the hypothesis, consider alternative explanations and make a logical connection between hypotheses, data and conclusions.
3. Your argument makes no sense. What is your evidence?	Partially	Asking for evidence is useful, but reviewer does not indicate which part of the paper is confusing them or what exactly they don't understand.
4. Your argument depends on weight being an inherited trait. What evidence do you have to support this assumption?	Yes	The reviewer has identified an assumption made by the writer and pointed out how the validity or invalidity of this assumption could impact the writer's conclusion.
5. Which of your hypotheses is best supported by the data?	Partially	The reviewer is specific in indicating that the writer did something well (posed multiple explanations) but provides only a vague indication that the writer needs to discuss the data more without indicating how or where they felt the writer's conclusions were lacking.

Appendix 2: **Criteria used in the Universal Rubric for Science Writing.** Contact the authors for more information or for a full version of the rubric including descriptions of performance levels (novice to proficient) for each criterion. A scoring guide including examples of student work at each performance level for each criterion is also available upon request.

Introduction	**Context** Demonstrates a clear understanding of the big picture; Why is this question important/ interesting in the field of biology?
	Accuracy and relevance Content knowledge is accurate, relevant and provides appropriate background for reader including defining critical terms.
Hypotheses	**Testable and consider alternatives** Hypotheses are clearly stated, testable and consider plausible alternative explanations
	Scientific merit Hypotheses have scientific merit.
Method s	**Controls and replication** Appropriate controls (including appropriate replication) are present and explained.
	Experimental design Experimental design is likely to produce salient and fruitful results (actually tests the hypotheses posed.)
Results	**Data selection** Data chosen are comprehensive, accurate and relevant.
	Data presentation Data are summarized in a logical format. Table or graph types are appropriate. Data are properly labeled including units. Graph axes are appropriately labeled and scaled and captions are informative and complete.
	Statistical analysis Statistical analysis is appropriate for hypotheses tested and appears correctly performed and interpreted with relevant values reported and explained.
Discussion	**Conclusions based on data selected** Conclusion is clearly and logically drawn from data provided. A logical chain of reasoning from hypothesis to data to conclusions is clearly and persuasively explained. Conflicting data, if present, are adequately addressed.
	Alternative explanations Alternative explanations (hypotheses) are considered and clearly eliminated by data in a persuasive discussion.
	Limitations of design Limitations of the data and/or experimental design and corresponding implications for data interpretation and conclusions are discussed.
	Implications of research Paper gives a clear indication of the implications and direction of the research in the future.
Use of Primary Literature	**Primary Literature** Writer provides a relevant and reasonably complete discussion of how this research project relates to others' work in the field (scientific context provided) using primary literature. Primary literature is defined as: • peer reviewed • reports original data • authors are the people who collected the data. • Journal produced by a non-commercial scientific association
Writing Quality	**Writing Quality** Grammar, word usage and organization facilitate the reader's understanding of the paper.

Notes and References

1. J. A. Shymansky, W. C. Kyle, Jr. and J. M. Alport, *Journal of Research in Science Teaching*, 1983, 20, 387-404.
2. C. M. Schroeder, T. P. Scott, H. Tolson, T.-Y. Huang and Y.-H. Lee, *Journal of Research in Science Teaching*, 2007, 44, 1436-1460.
3. K. Powell, *Nature*, 2003, 425, 234-236.
4. B. J. Alters and C. E. Nelson, *Evolution*, 2002, 56, 1891-1901.
5. K. Tobin, D. J. Tippins and A. J. Gallard, in *Handbook of research on science teaching*, ed. D. L. Gabel, Macmillan, New York, 1994, pp. 45-93.
6. R. Iyengar, M. A. Diverse-Pierluissi, S. L. Jenkins, A. M. Chan, L. A. Devi, E. A. Sobie, A. T. Ting and D. C. Weinstein, *Science*, 2008, 319, 1189-1190.
7. D. A. F. Haaga, *Teaching of Psychology*, 1993, 20, 28--32.
8. B. E. Timmerman, D. C. Strickland and S. M. Carstensen, *Integrative and Comparative Biology*, 2008, 48, 226-240.
9. N. J. Pelaez, *Advances in Physiology Education*, 2002, 26, 174-184.
10. K. J. Topping, E. F. Smith, I. Swanson and A. Elliot, *Assessment & Evaluation in Higher Education*, 2000, 25, 149-169.
11. T. M. Paulus, *Journal of Second Language Writing*, 1999, 8, 265-289.
12. A. Russel, in *Calibrated Peer Review Symposium*, National Science Foundation, Texas A&M, 2007.
13. K. Cho, C. D. Schunn and R. W. Wilson, *Journal of Educational Psychology*, 2006, 98, 891-901.
14. J. Hafner and P. Hafner, *International Journal of Science Education*, 2003, 25, 1509-1528.
15. K. Lundstrom and W. Baker, *Journal of Second Language Writing*, 2009, 18, 30-43.
16. K. Cho, C. D. Schunn and D. Charney, *Written Communication*, 2006, 23, 260-294.
17. M. M. Nelson and C. D. Schunn, *Instructional Science*, 2009, 37.
18. A. E. Lawson, B. Clark, E. Cramer-Meldrum, K. A. Falconer, J. M. Sequist and Y.-J. Kwon, *Journal of Research in Science Teaching*, 2000, 37, 81-101.
19. A. E. Lawson, *Journal of Research in Science Teaching*, 1978, 15, 11-24.