CREATING ACADEMIC COMMUNITY FOR FIRST-GENERATION COLLEGE STUDENTS
A Graduate Student Instructor Guidebook

Peer Review Improves Undergraduate Science Writing Skills

Debra Conte*
Quantitative and Systems Biology Program
Center for Research on Teaching Excellence Graduate Teaching Fellow

Inadequate writing skills are a common problem in science classes. Writing scientific papers require a different skill type than writing prose and often learning the terminology is similar to learning a second language. Research has shown that activities such as group projects and peer review are helpful in addressing this challenge in other contexts. This project will use three sections of upper division physiology university students; one section is the experimental group and the other two are control groups. All students write five lab reports with a partner during the semester, which are used to practice peer review. The class is trained on how to review a paper and a rubric in the form of a feedback worksheet is provided. The class anonymously exchanges papers and review prior to turning them in for grading. The reports with the attached feedback worksheet and additional comments from the instructor are returned to the students and common problems are discussed in class. The students also write three formal manuscripts in teams of four with students from all three sections. The first manuscript grade will be used as the baseline to measure improvement in the scores on the remaining two manuscripts. I hypothesize that the students participating in the peer reviewing influence their manuscript team and receive better outcomes on their grades. The results indicate improvement in the manuscript grades, suggesting that peer review has a positive effect in a scientific laboratory setting.

Sponsored by the Fund for the Improvement of Post-Secondary Education,
Future Faculty Grant **P116V090031**

Peer Review Improves Undergraduate Science Writing Skills

By Debra Conte

Undergraduate students generally perceive physical science learning as primarily memorizing details and theories along with conducting prescribed experiments with known outcomes. And often this is exactly what they experience, especially in introductory courses. Laboratory classes are commonly developed for the specific task of orienting students to laboratory equipment use and experimental process design. Students conduct controlled experiments with known outcomes, thereby reducing instructional variables. The emphasis, then, can be on the technical aspects of the experiment. This is an appropriate method to introduce the basic practice; however, at the upper-division level these basic skills should be mastered and enriched by critical thinking and communication opportunities. Unfortunately, most students progress through their college courses, even those for physical science majors, without ever being tasked to communicate their laboratory experience in a manner that demonstrates their understanding as well as their ability to apply their knowledge. Similar to the rote quality of laboratory experiments, most students learn scientific terminology from textbooks or lecture slides, without the authentic experience of communicating their findings in the established "language" of the scientific community. Most colleges and universities require all students to include some kind of freshman composition class that should help with communication skills. Unfortunately, the traditional expository skills learned at this level often do not directly translate into scientific writing, both because of instructional time gaps and content differences. This disconnection between composition and science curriculum is problematic, as language acquisition and communication strategies are fundamental to learning science.

Most of what we customarily call "knowledge" is language, which means that the key to understanding a subject is to understand its language (Postman & Weingartner, 1971). It is not sufficient to only memorize terminology, being able to critically organize information in a proper context is just as important. For example, a child learning to speak will learn the words and meanings; however, it takes training and practice to use the correct words in the correct context with grammatical accuracy. Similarly, there is more to scientific writing than just learning the terminology. A student must develop critical thinking skills to organize thoughts into a logical, sequential order and to describe them with specificity. Writing a research report is not the same as a scribing a narrative of observations and protocols; this form of professional writing is a carefully constructed argument that presents data and strong support for conclusions. You have to be able to explain your actions and data in a persuasive manner that leads the reader to arrive at the same conclusions as presented. This is a skill that must be developed over time and through practice. A student cannot communicate in science strictly by memorizing the methods of scientific writing, just as one cannot become an accomplished writer by simply memorizing grammar rules. To make this difficult transition into professional writing, students benefit from sharing writing strategies via peer review. In this paper I will evaluate the contribution of the peer review process on student's critical thinking skills with respect to scientific research writing in a college laboratory course.

One traditional way to gain scientific literacy is by reading scientific papers. This exercise is one method of gaining a deep understanding of how scientific information is conveyed through the medium of research articles (Levine, 2001). A published paper demonstrates the finished product of introducing an idea and presenting the arguments systematically to persuade your reader to accept the hypothesis as possible. Understanding both some of the content and the appropriate use of the language of science is an essential component on the path towards scientific literacy (Osborne, 2002). By viewing a student trying to learn scientific "language" as akin to learning a second language, those English as a Second Language teaching strategies become a potential resource. A study conducted by Kristi Lundstrom and Wendy Baker (2009) investigated how second-language learner writing improved from peer review, showing the value of giving and receiving feedback. Their results showed that all students writing improved, with the students giving the reviews experiencing the most improvement.

Inadequate writing skills are a common problem in the science courses I teach. Research has shown that activities such as group work and peer review are helpful in addressing this challenge in other contexts. Similarly, prior studies have suggested that peer review is a powerful tool for learning even for second-language learners. It has been suggested by scientists that science has its own language and needs to be learned in order to be successful at communicating science. So it seemed logical to incorporate peer reviews into my physiology laboratory class to better instruct students in a much-needed skill set. During the process of peer review, the students are working together at a different level than conducting experiments. This increased interaction seemed to increase student comfort with offering criticism and asking for advice, resulting in better team dynamics. The goal of this classroom project was to incorporate peer review exercises into the laboratory curriculum for senior physiology students, with applications to any college science classroom with low or high-stakes writing assignments.

The Project

The project was structured to include three sections of upper-division physiology classes; one section of 24 students as the experimental group (Ex) and the other two sections of 36 total students as a control group. The control groups did not participate in peer review of the laboratory reports and were under the supervision of a different teaching assistant (TA). All students wrote a required five lab reports with a partner from within their own laboratory sections during the semester as part of the normal curriculum. These reports were the target for in-class peer review activities. The Ex group used a code in place of their name to ensure anonymity and jointly reviewed a lab report written by a different group of two from within their section, who also remained anonymous, prior to turning them in for grading. The class received training on how to review a paper and a rubric in the form of a feedback worksheet (see Appendices 1, 2) that was attached to the reviewed report when it was complete. I graded the reports prior to looking at the peer review to prevent any bias and added my own comments. I also provided comments on the peer review worksheet if there was any inaccurate or inflammatory feedback given, which occurred occasionally in the first review cycle until the skill could be more explicitly addressed. The reports with the attached feedback worksheet were returned to the students. Common problems noted on the feedback worksheets were discussed with the class.

The students additionally wrote three manuscripts in teams of four, comprised of students from all three lab sections. The first manuscript grade served as the baseline to measure improvement in the scores of the remaining two manuscripts. I hypothesized that the students participating in peer review would influence their manuscript team and all students in the class would earn higher grades with each project.

Participation in the project was voluntary; all students in the section received a description of the project and how it might benefit them in terms of skills. Although there were no additional incentives such as extra points offered, all students agreed to participate. This class met once per week for a three-hour laboratory. Most weeks the planned activities did not require the entire scheduled time allotment, so this additional activity did not impact the overall schedule.

After receiving participation agreement from the students, we reviewed the required components of a lab report and discussed in detail how to evaluate each component per the supplied rubric. The required sections in the lab report are: Purpose, Hypothesis, Methods/Materials, Results, Discussion and Conclusion (see Appendix 2). The Purpose section was not given much weight in grading, as this was generally a restatement of what was provided in the lab protocols. Additionally, the Materials/Methods section was not weighted as heavily either since this was also given in the protocol. However, it was a difficult transition for the students; many lost points here as they were instructed to provide specific methods with the assumption that the reader would know the procedures without including every detail. Students needed to be able to differentiate between reporting methods without presenting a protocol. The rubric described what was expected from each section for their review and my grading standards (see Appendices 2, 3).

At the beginning of the section when a lab report was due, all reports were submitted and then randomly redistributed to each group, ensuring that no group received their own report back. The lab reports were generally two to three pages in length, so the review did not usually take more than fifteen minutes to accomplish. During this time the students seemed to be actively discussing the reports and each time the activity was repeated the breadth of the comments and criticisms improved. The reports were graded, the review sheets analyzed, returned to the students and some of the most common mistakes that were commented on or should have been commented on were discussed. At this time students could contribute any additional feedback.

The three manuscripts were written based on the three largest laboratory experiments. The teams had to choose a subject within the area of the experiment (i.e. exercise effects on heart rate or respiration during the cardio-vascular lab), select the relevant data from a pool, and conduct background searches to assist in writing the introduction and to support their hypothesis in the discussion. The manuscript includes a pertinent introduction, clear statement of a hypothesis, specific explanation of methods and statistics, clear presentation of only the relevant results with descriptive text and figures, a concise but complete discussion of the results relating it to the 'bigger picture,' and a conclusion that summarizes how the hypothesis was defended and the importance of the overall findings. In addition, the manuscript had to include an abstract and be formatted per the author's instruction for submission to a journal of their choice. (see Appendix 3)

The first manuscripts were randomly divided between the two TAs for grading, the alternate TA graded the second manuscript, and the third was randomly divided again. This was to ensure that no groups had all of their manuscripts graded by the same TA and thus equalize evaluations.

Assessment

The initial assessment of the study was based on my observation of (1.) individual report improvements with respect to structure and logic and (2.) feedback quality and specificity gains. In the beginning the reports were poorly formatted and inconsistent. Although students had been given specific instructions on how to present their information, the resulting quality varied considerably. As the students gained experience with giving and receiving feedback, the reports became less variable and closer to the expected format. The same trend was observed in the feedback quality. The students were given detailed instructions on how to evaluate the reports but initially there were only a few, general comments recorded on the feedback form. As the semester progressed and they were exposed to feedback on their own reports and manuscripts, the specificity and clarity of the comments seemed to improve. Because the other groups did not participate in the peer review activity and the other TA did not use the same grading rubric, or even the same sectional expectations, there was not enough similarity to compare the grades on the lab reports between the sections. However, because they interacted with the Ex students in writing the manuscripts and received intensive feedback from both of the TAs there can be an assumption of influence between the groups.

The true measure of the activity's effectiveness is the improvement of their manuscript grades. The level of writing and critical thinking required in the manuscript far exceeds the requirement for lab reports. The lab reports express findings that are prescribed; although this level of writing is relevant for practice, it does not require the same level of critical thinking expected in other forms of professional writing. The additional challenge for the project is that the manuscript is written in a larger group (4) and required more developed teamwork skills. All three sections were intermixed among the manuscript writing groups creating variations within each group as to the percentage of students trained in peer review and those who were not. Two groups did not have any Ex students; six groups had one Ex student, four groups were split evenly with two Ex students, two of the groups had three Ex students and only one group in which all four members were from the Ex section. This made it difficult to segregate effects from the training, but this first project had to remain within the constraints of an existing curriculum. Interestingly, the improvement of the group containing all Ex students over the three manuscripts was equal to the rest of their class members that were mixed with students from other sections, suggesting that there was some degree of peer influence between the manuscript team members that participated in the peer review and those that did not. The Ex group had higher average scores on each of the manuscripts vs. the control group, 71% vs. 67%, 77% vs. 74% and 88% vs. 85% (figure 1), however the control group experienced a 3.6% greater improvement in their grades overall compared to the experimental group (27.1% vs. 23.5%). One possible explanation is that the team members from the experimental group may have become more confident over time to influence their control group team members or perhaps taken a leading role in preparing the manuscripts. All team members receive the same grade on the manuscript regardless of who exerted more work or input. It is possible the students who felt they had more skill to achieve a higher grade for the group may have become dominant as the value of a high score on the manuscript became more critical towards their final grade. A possible cause of this disparity could be due to the lower beginning point for the control group. By the time the first manuscript was written, the Ex group had already had some peer review training and additional writing instruction that the control group did not receive. There was not sufficient time for the Ex group to have had much influence on the control groups yet. Surprisingly, when the groups were analyzed depending on the number of control members (figure 2) the two most notable data points are the groups with three control and zero control. The group with three control had the lowest starting scores and perhaps caused them to accept the Ex members input sooner. The group with no control students made the big jump on the second manuscript suggesting the training was being effective and they did not have to influence any control members. The data shows that with each manuscript the difference between the groups becomes closer, suggesting improved team dynamics.

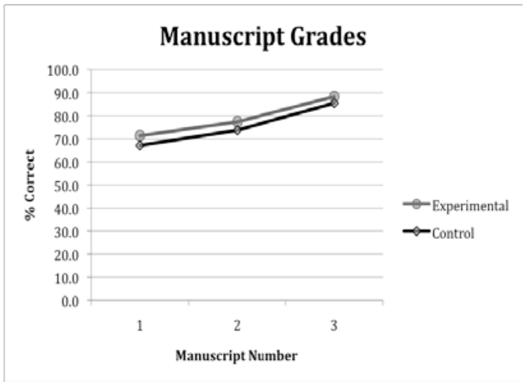

Figure 1: The experimental group was 29% higher than the control with the 1st manuscript but the difference was only 12% by the renal manuscript.

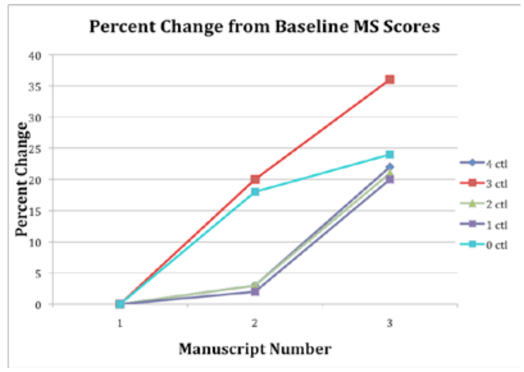

Figure 2: The group that had only 1 Exp member showed the largest improvement from baseline.

During the last class meeting, project participants were given a survey (figure 3) to evaluate how they felt about the project and the usefulness of the peer review activity. The response rate to the survey was nearly 100% with only one student not available to participate. To provide a little perspective on the group's prior writing experience, they were asked if they have previously taken a scientific writing class and 98% of the students had a positive response and 74% of the students had previously participated in peer review (figure 3). The third and fourth question on the survey asked the students how comfortable they were with giving and receiving feedback; 78%-87% were comfortable or very comfortable with both giving and receiving, however most responders were definitely more comfortable with receiving the feedback. More than half of the students felt peer review helped to improve their skills moderately to very much and only a small percentage felt that it had no impact at all.

The survey collected feedback regarding the effectiveness of improving each section of the report, with the exception of the methods section, which was split somewhat evenly. The moderate to very helpful responses were much higher (65%) than the very little or not at all responses (33%). Since this survey was given to them prior to the final manuscript grade being released, this clearly suggests that the students perceived the peer review activity as helpful regardless of their grade improvement.

	Question	Yes	Very Uncomfortable (%)	Uncomfortable (%)	Comfortable (%)	Very Comfortable (%)
1	Have you taken a scientific writing class	96%				
2	participated in peer review of written materials?	74%				
3	How comfortable do you feel giving feed back to peers?		9	4	57	30
4	How comfortable do you feel receiving feedback from peers?		22		30	48
			Not at all (%)	Very Little (%)	Moderately (%)	Very much(%)
5	How much do you feel that peer review has helped improve formatting your scientific papers properly?		4	26	61	9
6	How much do you feel that peer review has helped improve writing your hypothesis properly?		9	30	48	13
7	How much do you feel that peer review has helped iimprove descibing your methods properly?		14	23	50	1
8	How much do you feel that peer review has helped improve stating your results and using figures properly?		4	22	52	22
9	How much do you feel that peer review has helped improve developing a discussion to support your hypothesis?		4	26	57	13
10	How much do you feel that peer review has helped improve your writing an impactful conclusion?		5	32	55	9
11	Do you feel practicing peer review has increased your ability to evaluate reports for logical structure and content?	82				

Figure 3: Results from the final survey reflect most students perceived the activity to be beneficial.

Conclusions

Persuasive, professional written communication is an expectation of the scientific community. If we cannot share our discoveries with others, our knowledge and technology cannot progress. These skills are beyond the written word but include, "an integration of words, diagrams, pictures, graphs, maps, equations, tables, charts, and other forms of visual mathematical expression" (Lemki, 1998). These skills need to be practiced as well as taught, which can be challenging since the writing methodology is dissimilar to standard expository writing instruction in freshman year. Because this transition to scientific writing is so dramatic, student challenges are similar to learning a second language. Applying some of the effective teaching strategies for second language learners has positive impact on writing-intensive science classrooms, like mine, as well.

This project focused on the specific goal of improving the writing skills of upper-division college students after historical observation of a weakness in this area. Peer review has been shown to improve student's critical thinking (Todd & Hudson, 2007) and has been effectively used in second language learning

(Lundstrom & Baker, 2009; Swain, Brooks, & Tocalli-Beller, 2002). Like most graduate student instructors in the sciences, I was not the instructor of record for the course. With my project operating within an established course structure, it was difficult to control some variables; however, the peer review student learning outcome results indicate that not only did those who participated directly in the project benefit, those students who had contact with the experimental group in other activities also benefited.

The peer review exercise was performed using less significant laboratory reports because these reports were written entirely by students in the experimental group. This was the only material they produced independently, providing an element of internal control as well as adding five additional opportunities to practice. In the context of this review the results of their lab report scores (not presented here) showed an improvement over time as well as the quality of the feedback critiques. During the first attempt at this activity many of the reviewers simply circled the number that matched their evaluation of that section making only a few general comments such as "Need to check spelling" or "Graphs are hard to read." After receiving critiques on both lab reports and manuscripts, the comments became more critical, "Discussion doesn't apply to hypothesis" or "In the figures [you need to] state what you are measuring, is it percent mean?" This improved feedback quality suggests that students were better at understanding what was expected in each of the report sections and evaluating whether the information presented made sense. This improvement also implies a more focused interpretation of what was presented by their peers. I did not have a quantitative way to measure the quality of the feedback; however, input from the students and my personal observation support this conclusion.

The student surveys supported a perceived improvement from them, even though they had not received final manuscript grades. Their final lab report quality improved and their group interaction was enhanced by peer review. Several students mentioned that because of peer review, they felt more comfortable approaching the more successful groups to ask for advice. The general atmosphere in the class was more collegial and interactive than my prior experiences with this course or what was reported in other sections. Although these are not tangible results, anecdotal evidence indicates that peer review supported a more productive learning environment than otherwise. This improved atmosphere is important, as group work is fundamental as well to success in the sciences.

Students were asked in a survey to provide comments and recommend improvements. Most students felt the process was important enough to improve and repeat. Some suggested improvements included having all of the sections participate in peer review, including a better example of the 'perfect' paper to emulate, and providing the feedback sooner. One improvement I would make is to discuss the common problems found during the peer review, immediately during the same class while the information is fresh in their minds and to allow time to incorporate the lessons learned into practice quicker.

In conclusion, the experiment was successful at showing the value of peer review in a scientific context. The results supported my hypothesis that training even a portion of the students would have an overall influence on the entire class. The final results could be substantial if this teaching strategy (peer review) could be incorporated in some form in all scientific laboratory classes.

References

Lemki, J. L. (Producer). (1998, 05/10/2010) Teaching All the Languages of Science: words, symbols, images and actions. *Barcelona Papers*. retrieved from http://academic.brooklyn.cuny.edu/education/jlemke/papers/barcelon.htm

Levine, E. (2001). Reading Your Way to Scientific Literacy. [Guides-classroom]. *Journal Of College Science Teaching, 31*(2), 122-125.

Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. [doi: DOI: 10.1016/j.jslw.2008.06.002]. *Journal of Second Language Writing, 18*(1), 30-43.

Osborne, J. (2002). Science Without Literacy: a ship without a sail? *Cambridge Journal of Education, 32*, 203-218.

Postman, N., & Weingartner, C. (1971). *Teaching as a Subversive Activity*: Delta.

Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). Peer-Peer Dialogue as a Means of Second Language Learning *Annual Review of Applied Linguistics* (pp. 171-185): Cambridge University Press.

Todd, V., & Hudson, J. C. (2007). Using Graded Peer Evaluation To Improve Students' Writing Skills, Critical Thinking Ability, And Comprehension Of Material In A Principles Of Public Relations Course. *Journal of College Teaching & Learning, 4*(10), 39-46.

Appendix 1

Author I.D. #'s:____________________ Reviewer's I.D. #'s________________

Peer Review Worksheet

0 = N/A, 1=Very low quality, 2=Low quality, 3= High quality, 4=Very High quality

1. Proper format, including: Purpose, Hypothesis, Materials/Methods, Results, Discussion, Conclusion.

   0 1 2 3 4

   Comments:

2. Clearly stated hypothesis.

   0 1 2 3 4
   Comments:

3. Clearly presented results including graphical illustration with descriptive figure legend. Includes descriptive text describing results without interpretation.

   0 1 2 3 4

   Comments:

4. Discussion effectively applies results to prove or disprove hypothesis. Relates data into physiological context.

   0 1 2 3 4

   Comments:

5. Presents clear conclusion based on results and discussion.

   0 1 2 3 4

   Comments:

6. Uses appropriate scientific language in a logical, easy to follow order.

   0 1 2 3 4
   Comments:

Appendix 2

BIO 161: Lab Report Evaluation Form

Total points available 25

1) Hypothesis (2 pts.)

Is hypothesis clearly stated and specific?

2) Methods. (2pts.)

Is sufficient information presented to clearly understand how the study was conducted? Formatted correctly?

3) Results. (5 pts.)

(a.) Is description of results presently clearly?
(b.) Are figures correct?
(c.) Are tables correct?
(d.) Is sufficient and the appropriate data presented to address the hypothesis?
(e.) Formatted correctly?

4) Discussion. (10 pts.)

(a.) Does the discussion follow a logical order?
(b.) Do discussion topics relate to the hypothesis?
(c.) Does the discussion connect the lab to human physiology?
(d.) Is the study summarized in the concluding paragraph?

5) Conclusion. (3 pts.)

(a.) Are the conclusions logical and supported by the data?
(b.) Is there a "take-home" message?

6) Overall. (3 pts.)

(a.) Are grammatical errors minimized? (spelling, sentence structure etc)
(b.) Is it well written and easy to follow?
(c.) Proper report format?
(d.) Are all components present? (Title, purpose, hypothesis/research question, methods/materials, results, discussion and conclusion)

Appendix 3

BIO 161: Manuscript Evaluation Form

Group Members: Score: /50

1) Title. (1 pts.)

Effective & sensible?

2) Abstract. (7 pts.)

(a.) Well written?

(b.) All components present? (introductory sentence, hypothesis/research question, methods, results, conclusions, implications)

3) Introduction. (8 pts.)

(a.) Hypothesis clear and presented well?

(b.) Is background information sufficient to support pretense for study?

(c.) Is information cited correctly?

4) Methods. (5 pts.)

Is sufficient information presented to clearly understand how the study was conducted? Formatted correctly?

5) Results. (10 pts.)

(a.) Is description of results presently clearly?

(b.) Are figures correct?

(c.) Are tables correct?

(d.) Is sufficient and the appropriate data presented to address the hypothesis?

(e.) Formatted correctly?

6) Discussion. (12 pts.)

(a.) Does the discussion follow a logical order?

(b.) Do discussion topics relate to the hypothesis?

(c.) Are the conclusions logical and supported by the data?

(d.) Are implications of the study presented?

(e.) Is the study summarized in the concluding paragraph?

(f.) Is there a "take-home" message?

7) References. (5 pts.)

(a.) Are there at least 10 peer-reviewed citations?

(b.) Are they used correctly?

(c.) Are they formatted correctly?

(d.) Are they cited within the text of the paper?

8) Overall. (2 pts.)

(a.) Are grammatical errors minimized? (spelling, sentence structure etc)

(b.) Is it well written and easy to follow?

(c.) Is it formatted like a scientific manuscript?