

THE IMPACT OF PEER REVIEW ON CONSTRUCTING ARGUMENTS BASED ON THE
CLAIM-EVIDENCE-REASONING FRAMEWORK

by

Kevin Knapik

A professional paper submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Science Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2018

©COPYRIGHT

by

Kevin Michael Knapik

2018

All Rights Reserved

TABLE OF CONTENTS

1. INTRODUCTION AND BACKGROUND	1
2. CONCEPTUAL FRAMEWORK	3
3. METHODOLOGY	10
4. DATA AND ANALYSIS	15
5. INTERPRETATION AND CONCLUSION	33
6. VALUE	36
REFERENCES CITED	42
APPENDICES	44
APPENDIX A Experimental Design Performance Assessment.....	45
APPENDIX B Rubric for Experimental Design Performance Assessment	48
APPENDIX C Example of an Experimental Design Evaluation Exit Slip	51
APPENDIX D Peer Review Likert Survey	53
APPENDIX E Student Interview Questions.....	56

LIST OF TABLES

1. Data Triangulation Matrix.....14

LIST OF FIGURES

1. Student Agreement on Experimental Design Evaluation Exit Slip	16
2. Pre- and Post-Peer Review Attitude Survey Question One Results	20
3. Pre- and Post-Peer Review Attitude Survey Question Two Results.....	22
4. Pre- and Post-Peer Review Attitude Survey Question Three Results.....	23
5. Pre- and Post-Peer Review Attitude Survey Question Four Results	25
6. Pre- and Post-Peer Review Attitude Survey Question Five Results.....	26
7. Pre- and Post-Peer Review Attitude Survey Question Six Results.....	27
8. Pre- and Post-Peer Review Attitude Survey Question Seven Results	27
9. Pre- and Post-Peer Review Attitude Survey Question Eight Results	28
10. Pre- and Post-Peer Review Attitude Survey Question Nine Results	29
11. Pre- and Post-Peer Review Attitude Survey Question Ten Results.....	30
12. Student Dependent Variable on Experimental Design Performance Assessment	31

ABSTRACT

Classroom peer review strategies have the potential to help students engage in a vital practice of the scientific community. Performing peer and self-critique helped students calibrate their personal level of skepticism so that they can accept accurate claims more frequently when the claim is paired with data. Furthermore, students hold themselves to a higher standard when selecting data to be gathered in a designed experiment after applying critique to multiple approaches to the same assignment. This study was implemented to determine whether explicit peer review and self-reflection strategies impact student perception of the peer review process. Additionally, the study investigated whether the same strategies impact student success on performance assessments where they design their own experiment and critique a fictional experiment. The results indicated that practicing peer critique establishes a sense of confidence in the process. Students thought that peer review was valuable in making outcomes more accurate. Moreover, students were able to narrow their dependent variables more effectively when designing their own experiment and they were more effective at identifying elements of the fictional experiment that were problematic.

INTRODUCTION AND BACKGROUND

Evidence-based argumentation is an essential skill for a student hoping to engage in productive scientific discourse. Furthermore, selecting appropriate variables to measure is the first step in developing a defense of claims made at the conclusion of an experiment. In science education, there is an expectation that students are being prepared to interpret scientific findings and to be ready, students must be supported in argument skill development. The claim-evidence-reasoning (CER) framework for scientific argumentation has the potential to support students in developing essential argumentation skills. In this framework, students are given a format to follow that includes making a claim, supporting it with evidence, and explaining how the evidence relates to the claim. An important part of this framework is selecting appropriate and sufficient evidence; however, what appears to be appropriate and sufficient to one student may not be to another. To find new ways to enhance the experience of practicing scientific argumentation, I rediscovered the potential that is inherent in the peer-editing process.

I chose to incorporate purposeful peer-editing in my Honors Biology course at Evergreen Park High School in Evergreen Park, Illinois. Of the twenty-one students in the class, two were African American (10%), four were Latino (19%), and fifteen were white (71%). The demographics of the school as a whole are different, with white students making up 48.9%, African American students making up 34.7%, and Hispanic students making up 12.9% of the student body. Evergreen Park Community High School (EPCHS) is relatively small for the Chicago suburbs, standing at around 850 students with 28.4% of the student body coming from low-income households (Illinois State

Board of Education, 2016). The climate of EPCHS mirrors the community of Evergreen Park in that it is small enough for all of the students to know nearly every peer in the school. EPCHS competes with several private schools in the area for students. Evergreen Park is located just outside of the south side of Chicago. The students in Honors Biology are freshmen who scored exceptionally high on the science and math section of their entrance exam or sophomores who demonstrated above-average ability in their non-honors freshmen science course.

I chose my Honors Biology course to implement purposeful peer-editing in an effort to determine the impact of the technique on student argumentation skills. I selected this course because I have been teaching biology for five years and the other two courses that I teach are still in their early stages of development. The students often arrive at my classroom with limited exposure to argumentation strategies and they report having little experience with editing peer work in science. As part of our curriculum, students frequently design, present, and, in some cases, implement their own scientific experiments. During presentations, students are expected to ask questions to help clarify the experiment that was just presented to the class. In the past, this process served as a form of peer-editing because the presenters were often forced to confront mistakes in their experimental design. During the presentation of their work, students were expected to defend their choice of dependent variables. Selecting a precise dependent variable is an essential aspect of the CER framework of scientific argumentation and purposeful peer evaluation has the potential to improve overall student choice and defense of evidence.

As a science department, we decided to use aspects of the American Modeling Teachers Association (AMTA) curriculum. I supplement the materials found in the AMTA curriculum with some lessons from Project Novel Education for Understanding Research on Neuroscience (Project NEURON) and Progressing through the Ages: Global Change, Evolution, and Societal Well-being (PAGES), based out of the University of Illinois, which employ similar student modeling techniques found in the AMTA lessons. One goal of both programs is to help students develop the skills necessary for participation in science and engineering fields, such as evidence-based argumentation (Project NEURON, 2015). Portions of the Project NEURON lessons and the AMTA lessons have been woven together to build an understanding of the nature of modern scientific practices. Participation in these programs lead to the generation of my focus question, How does peer evaluation impact student performance on evidence-based argumentation assessments? Additionally, this project includes two sub-questions. 1. How does the inclusion of peer-editing strategies impact thinking about the scientific process? 2. What changes in student evidence selection resulted from engaging in the peer evaluation process?

CONCEPTUAL FRAMEWORK

Strong student performance on argumentation style science assessments is often characterized by the use of appropriate and sufficient evidence and an explanation for how this evidence supports claims that are made about a phenomenon. The claim evidence reasoning framework was developed to support students in constructing strong verbal and written arguments based on scientific principles. In this framework, a claim is

the student's initial answer to the question, pieces of evidence are facts that are rooted in scientific data, and reasoning is the explanation that the student uses to connect the data to the claim. This framework gives students a structure on which to base their written and verbal communication in the science classroom (McNeill & Martin, 2011).

Scientific argumentation is identified as an area in which students need further support, even at the undergraduate level. Professors in undergraduate programs expect that students are able to successfully express themselves by connecting theory to evidence while also considering alternative explanations as part of a rebuttal. Researchers in one such study found that students often progressed in making simple arguments by taking theory and applying it to evidence to support a claim, but students did not make the same gains in identifying alternative explanations and developing a rebuttal during their undergraduate career. Students that are unable to develop higher level arguments often make errors in evidence selection (Shen, 2013).

Teachers at the high school level are found to be ill-equipped to support their students to meet the needs identified at the collegiate level and in their own classroom. Teachers using the claim evidence reasoning framework tend to accept incomplete claims that do not include generalization of the claim. Generalization is extending explanations beyond the direct circumstances of the investigation and students who apply generalization to their arguments further explain their claims to include aspects of the scientific phenomenon outside of what they observe directly in lab conditions. Teachers that omit generalization from their formative student support might accept a direct answer about the lab that does not include how the circumstances of the lab apply to the bigger

model of the phenomenon (Shemwell et al., 2015). Ignoring a lack of generalization in student answers may limit a student's ability to meet the standard set at the collegiate level, where students are expected to consider more than just their own perspective.

With a framework in place, teachers and students can be supported in their growth and development in the area of argumentation. Researchers from one study found that while teachers rarely identify generalization as an important component of scientific argumentation in initial surveys, teachers increase their ability to construct generalized claims and recognize the value of generalization after using formative tools that ask the user to take a claim and apply it to similar situations (Shemwell, Gwarjanski, Capps, Avargil, & Meyer, 2015). Regardless of whether teachers recognize that there is need for argumentation support in their classroom or not, the researchers posited that teachers may not be fully prepared to coherently use tools associated with scientific argumentation that are meant to help students meet the standard set by tertiary educators and beyond (Shemwell et al., 2015).

Questioning strategies employed by teachers impact student development of skills in applying generalization to claims, evidence selection, and identifying alternative responses to the same problem. Teachers commonly use one major strategy for asking questions in the classroom. Researchers in one study identified this as the known answer strategy where there is a specific answer that the teacher indicates as the one correct answer. This same study found that more open-ended questions resulted in more student-student interaction. Ultimately, the open conversations were correlated with students in these classrooms using more evidence and reasoning in their verbal and written

argumentation (McNeill & Pimentel, 2009). The researchers in this case identified that different questioning strategies might be used when the goal of a discussion is different. When introducing a topic, traditional known-answer questions might be an effective tool but when supporting students in developing argumentation strategies is the goal, open ended questions can be used to probe for student ideas and increase exposure to competing ideas.

In an effort to enhance student argumentation in science and student willingness to engage in science inquiry, teachers are turning to a variety of techniques including having their class generate models. Coherency between activities is an integral component of a course that encourages student model building. Authentic modeling can be defined as using activities to gather information about a phenomenon with the intention of refining a cohesive explanation, while the activities also demonstrate real world application (Gjalt, Bulte, & Pilot, 2016). This can be established through continual student reflection on the activities. If the framework for the course includes adequate time for reflection, students can draw connections between different activities that establish clarity for different pieces of the model. Gjalt et al. (2016) identified difficulties in establishing built in reflection time. Difficulties persisted when students maintained a view of science where information is accepted when it comes from books, teachers, or so-called experts. Based on these findings the researchers recommend that students be provided a scaffolded framework to establish greater value in the reflective process.

Student engagement in modeling is not dependent on pre-existing content knowledge. With the proper support, high school students can become more proficient in

generating simplified explanations that incorporate information from several sources, including inquiry labs and activities. Researchers in one study found that there was little difference between students with minimal experience with content and those with more background knowledge in generating strong arguments after engaging in a modeling curriculum (Fortus, Schwartz, & Rosenfeld, 2015). Fortus et al. (2015) indicated that having students engage in developing models can be beneficial in helping students develop content knowledge but it can also be beneficial to teachers who establish model building skills as a goal of the classroom. This can be helpful to students regardless of their pre-existing exposure to model building.

Modeling takes many different forms but usually involves students gathering information, making a product, and evaluating their end result. The models are usually based on physical experimentation or performance simulation, but the final product can be mathematical, graphical, physical, or virtual. There is continued discourse over which type of model is most effective in helping students gain a better understanding of the material. Researchers in one study found that a combination of physical and virtual models was more effective than either modality on its own in helping students confront misconceptions about the material (Blikstein, Fuhrmann, & Salehi, 2016). In their framework, students were asked to perform a physical experiment, develop a virtual model, detect discrepancies, reflect on the reason for the discrepancies, and remedy the problem in the virtual model. Researchers in the study found that combining virtual and physical models and evaluating their inconsistencies is an inherently reflective practice. Students of this study demonstrated gains in understanding both the micro and macro

nature of the content being taught. Students also demonstrated progress in taking a complex system and breaking it down into micro level rules (Blikstein et al., 2016).

Reflection and evaluation are crucial aspects of the science classroom. The Argument-Driven Inquiry Instructional framework is one model used to create an environment in which the students can practice self-reflection, peer evaluation, and revision. In this framework, students begin by identifying the task and the overarching research question. They then develop their method for collecting and analyzing data, with the purpose of developing a tentative argument. The various groups of students then engage in an argumentation session and if needed, groups can develop new ways of collecting additional data. Lastly, the report is written and submitted for peer review and revisions are made based on peer review. Researchers from one study of undergraduate biology students found that when instructors use the Argument-Driven Inquiry methods, students can develop higher level writing skills regardless of their level upon entering the course. They argued that this method was a practical way to provide more science writing practice by transferring the responsibility of evaluation to the students (Walker & Sampson, 2013). The peer evaluation process sets up a release of responsibility to the students where they take on a task typically performed by the instructor. This transfer allows students to practice for the rigorous review process that exists in the scientific community in a controlled setting where they get feedback on how they critique their peers. Furthermore, students have more exposure to argumentative writing and how other students progress in this writing style.

Engagement in peer review is associated with gains made in student attitude toward their own abilities to digest scientific literature and critically reflect. Students who have a structure for peer review report greater confidence in their abilities to read and offer suggestions on papers that they read. In one study where students self-reported feelings about the peer review and science writing process, researchers found that students were likely to report a greater appreciation toward peer review after engaging in a course that had an established format of peer review. Students in this study reported that the final peer review session served as the greatest determining factor in promoting their understanding of the science content. Students reported that the practices of engaging in peer review and evaluating the data section in papers were chief determining factors in strengthening their science literacy (Geithner & Pollastro, 2016). Students found the process rewarding and reported that they could see value in the process of peer review as it pertained to future careers.

Peer review strategies are being employed by tertiary educators that are making the transition to inquiry-based coursework and there is evidence to suggest that peer review is effective in supporting students as they develop science writing skills. In one study of undergraduate participants, researchers found that student gains depend on the level of guidance provided to a reviewer through implementation of a rubric. Researchers in this study compared student reviewers who used an experimental rubric that asked them specific follow up questions after assigning a grade for each category of the paper to student reviewers using a control rubric that simply asked them to explain the grade for each section. Researchers found that when provided with the experimental rubric,

reviewers were able to provide more meaningful feedback that could be used to improve student outcomes in writing (Kelly, 2015). This result demonstrates that peer review strategies differ in their outcomes and this ultimately depends on the level of students using the strategies. In upper level coursework, where writing is more intensive and inquiry is expected, teachers are preparing their students for a more direct transition to the peer review model employed by the scientific community. Investigators in one study that observed students in higher level coursework where the students transitioned from a rubric based peer assessments to free-format peer review. Investigators found that the students accepted and desired the peer review intensive curriculum according to data collected after completion of the coursework (Glaser, 2014). Students making the transition to college coursework will be expected to effectively engage and develop their skills in providing feedback to peers.

METHODOLOGY

The focus of this action research was to determine the impact of integrated peer review strategies on student success in developing evidence-based arguments and on student attitudes towards science. Twenty-one students in one Honors Biology classroom participated in this study across two units. All students in this section of Honors Biology were in their first year of high school and this was their first science course at Evergreen Park Community High School (EPCHS). EPCHS competes with several private schools in the area for students. Evergreen Park is located just outside of the south side of Chicago. The students in Honors Biology were freshmen who scored exceptionally high

on the science and math section of their entrance exam or sophomores who demonstrated above-average ability in their non-honors freshmen science course.

Self-reflection and peer review strategies were added to one Honors Biology section as part of the routine for the unit. This group of students typically achieve at a high level academically and most will continue taking science classes beyond the three-credit requirement. The science department at EPCHS uses pieces of the American Modeling Teachers Association (AMTA) curriculum. In Honors Biology, the materials found in the AMTA curriculum are supplemented with lessons from Project NEURON and PAGES, based out of the University of Illinois, which employ the same student modeling techniques found in the AMTA lessons.

Students were given Pre- and Post-Treatment versions of the Experimental Design Performance Assessment to determine their ability to design and defend experiments that help answer a purposefully vague experimental question (Appendix A). The pre-test was given immediately before participating in the treatment unit and the post-test was given on the last day of the unit. In this assessment, students needed to draw a model of a controlled experiment that would draw a conclusion to support an answer to the experimental question. Students were assessed on their ability to select an appropriate and measurable dependent variable. A rubric for the Experimental Design Performance Assessment (Appendix B) was used to grade students on whether all components of a controlled experiment were included and whether the dependent variable could be used to support a conclusion to the experimental question. The students were given a post-test with minimal differences that assessed the same skill at the end of the treatment unit.

Students were also given the initial Experimental Design Evaluation Exit Slip before starting the treatment unit (Appendix C). Versions of the exit slip that were given during the treatment unit were peer reviewed by classmates. Student reviewers were instructed to provide written questions that would help the writer construct a more precise answer. Students were given the final version of the Experimental Design Evaluation Exit Slip at the end of the treatment unit to determine their ability to explain criticism of the fictional experiment.

Students were given the Peer Review Likert Survey at the beginning of the treatment unit to determine their experience with peer review strategies (Appendix D) prior to the treatment. The questions gauged how much students used peer review and how valuable they considered different components of the peer review process. Furthermore, students reported to what extent they agreed that they felt they could construct scientific arguments. At the end of the treatment unit, students were given the same Likert survey to determine whether students perceived that their attitudes changed after incorporating peer review strategies into their daily routine.

During the treatment unit, students were exposed to several self-review and peer review strategies. The first of these strategies was student self-reflection. At the beginning of each period, students wrote a reflection about what they did in class the previous day. Students shared what they wrote with the peers at their table. The second strategy that was introduced occurred after each major activity. Students completed exit slips that contained questions about the conclusions that could be drawn from the lesson materials. The students were instructed to explain how they knew that their answer was

correct. After completing the exit slip, the students then exchanged their papers with a classmate, who was instructed to circle a portion of the answer and ask a question about the original student's answer. Upon returning the paper, the original student wrote more to clarify the reviewing student's question. Lastly, students presented group work to the class and the audience was able to ask questions to help the presenting students come to a better conclusion. Prior to the presentations, students developed whiteboard models that included visual aids for the students to reference while they were presenting. In some cases, students were asked to write questions on Post-It notes to help groups improve their whiteboards before presentations. Student groups were given time to make additions to their whiteboards after students walked around the room posting questions on boards. All of these strategies were intended to help students consider how they might improve.

At the beginning and end of the treatment unit, four students were selected at random to complete an interview (Appendix E). The questions used in the interview were selected to determine how often and to what extent students used peer review in other classes. The interview participants also shared how valuable the students found the experiences. The interview participants received the same questions before and after the treatment unit with slight modifications. In the post-interview, students were told which strategies were part of the peer review process. The student responses were compared to determine whether their attitudes changed after completing the treatment unit.

The Data Triangulation Matrix contains the data collection strategies for the research period (Table 1).

Table 1
Data Triangulation Matrix

Focus Questions	Data Source 1	Data Source 2	Data Source 3
<p><i>Primary Question:</i></p> <p>What impact do peer review strategies have on students' abilities to construct arguments based on the claim evidence reasoning framework?</p>	Experimental Design Performance Assessment	Experimental Design Evaluation Exit Slips	Student Work
<p><i>Secondary Question:</i></p> <p>What impact does integration of peer review strategies have on student perceptions of the value of peer review?</p>	Peer Review Likert Survey	Student Work	Student Interviews
<p><i>Secondary Question:</i></p> <p>What impact does integration of peer review strategies have on selecting measureable dependent variables?</p>	Experimental Design Performance Assessment	Experimental Design Evaluation Exit Slips	Student Work
<p><i>Secondary Question:</i></p> <p>What impact does integration of peer review strategies have on student perception of their own ability to provide meaningful feedback?</p>	Peer Review Likert Survey	Student Interviews	Student Work

DATA AND ANALYSIS

After a two-week case study unit on the rise in Lyme disease due to forest fragmentation, students completed the Experimental Design Evaluation Exit Slip (Appendix C). The two-week unit included ten total lessons with eighty three minutes of instructional time. The unit ended with students performing choice experiments with disease free ticks in the classroom. In the exit slip, students chose to agree or disagree with the conclusion drawn by a fictional scientist. The scientist in the scenario just completed a choice experiment and used evidence to support the claim that ticks identify their host using heat detection. By design, the experiment in the problem included several factors that made it difficult to use the data generated as support for the conclusion. Twenty-one students completed the exit slip ($N=21$).

Prior to completing the treatment unit, six students (29%) did not agree that the scientist's conclusion was valid, with each student citing one of several reasons as to why the experimental design made it difficult to draw a conclusion. One student (5%) agreed with the scientist with some noted reservations and fourteen students (67%) agreed with the fictional scientist about the conclusion that was stated in the problem without reservation. After completing the treatment unit, eleven students (52%) chose to disagree with the conclusion drawn by the fictional scientist, five students (24%) said they could agree with certain modifications made to the experimental design, and five students (24%) chose to agree with the conclusion without reservation.

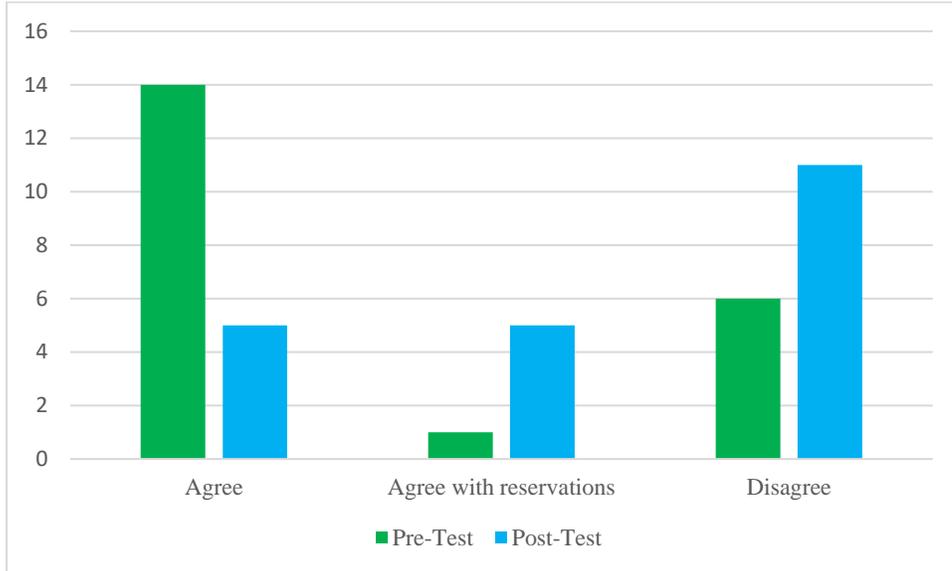


Figure 1: Number of students that chose to agree with no reservations, agree with noted reservations, or disagree with the final conclusion of the fictional scientist, ($N=21$).

There were five students that agreed with the scientist's conclusion with reservation and these students varied in their reasons for making their selection. In support for her decision after the treatment unit, one student said "I agree for this trial but this scientist should do multiple tests in order to have evidence to support his or her claim." This demonstrates that the student considered the number of trials as potentially problematic in determining the sufficiency of evidence in supporting the claim. Two other students recommended a similar modification to the experiment. A fourth student in this group said "According to the data she collected, it would appear that her conclusion is valid. However, because she only tested one possible way, we cannot say for sure that this is the only way. They (ticks) may favor a warmer climate but this does not mean that they sense heat to find their hosts." This student is suggesting that because the scientist was only doing a trial with the heat lamp as the experimental choice, she is not eliminating other possibilities that we observed during the choice box experiment, such

as carbon dioxide and vibration. The students that agreed with the scientist's conclusion with suggested modifications demonstrated that they were capable of communicating their skepticism of an imperfect experimental design and their reasoning for being critical.

The students that agreed with the conclusion drawn by the scientist without reservation had similar reasons for agreeing or were unable to communicate their reasoning for why they agreed. One student in this group said, "I agree because when they use no heat, they do not move towards or away from the lamp and they move nowhere so obviously where they go, depends on heat." This student was able to find patterns in the data but she was unable to communicate what the patterns meant to her and how they connected to her decision to agree. A second student in this group said, "Yeah because more went toward the heat lamp than ones who didn't." This student used evidence to support his argument but did not expand on why this experimental design was valid. Lastly, one student in the group that agreed with the conclusion supported his argument with evidence not found in the scenario. This student said, "I do agree with this conclusion because I think it is enough data to make a conclusion about this. Also, I agree with this also because it is correct based off our experiments." This student had preconceptions based on our discussion before doing the choice experiments. Regardless of their reason for agreement, students in this group were ineffective in connecting the data and components of the experimental design to explain why they agreed that the scientist's conclusion was valid.

The students that disagreed with the conclusion drawn by the fictional scientist varied in their reasons for disagreeing. Nine students (82%) disagreed with the conclusion because they did not consider the evidence to be sufficient to make the statement made by the scientist. One student in this group said, “I think you need three trials plus one control round before making any kind of conclusion.” This student identified that she may be able to agree if the data we were using were an average of many trials. A second student in this group said, “I cannot agree with the scientist because they only did one trial, and the amount that went towards the heat was only half of a small number.” This student explained that the sample size is problematic when considering just one trial. Two of the students that disagreed (18%) disagreed because the environment difference between each side might not be adequately controlled. One of these students said, “They (the scientist) need to know whether the heat lamp is giving off anything else. Overall, the 1st trial varies too much to determine whether they can detect heat.” This student suggests that the heat lamp may be giving off a second factor, such as light, that the ticks may be detecting. Students that disagreed with the conclusion demonstrated a high level of skepticism when critiquing conclusions but they varied in their abilities to extend their answers to explain why they disagreed with the conclusion of the experiment.

Throughout the treatment unit, students responded to self-reflection prompts as a component of their student work. The students varied in their responses but several made connections between peer review and getting better at doing science. On the sixth day of the unit, students responded to question, “How can you get better at doing science?” In

response to this question, eight students identified that they should be asking more questions. One student that fell into this group wrote, “I can get better at science by improving the way I approach problems. This consists of asking questions and focusing all the way and staying on topic.” This student identifies asking questions as a skill that someone can improve. In response to the same question, another student wrote,

You can train yourself to get better at the concept of doing science by asking questions about things you may not have thought about before. You can create experiments to learn more about new things. Trial, error, and other people interested in science are things that can help you get better. Trial and error help[s] you get better in the experimenting process and other people can poke holes in your experiments to help you become more thorough.

This student specifically identified peer critique as valuable in the improvement process. Additional students indicated that science requires expanding their minds to consider outcomes that they do not identify immediately. Responses to this item of student work were helpful in determining whether students were considering the role of critique in getting better at finding the most defensible argument about a given problem while constructing group arguments.

The Peer Review Attitude Survey was used to collect data on perception of the role of peer review in science classrooms and how it was used. A Wilcoxon Rank Sum Test was used to determine whether the results were statistically significant at $p < 0.05$. The data generated during the Peer Review Attitude Survey was ordinal and a Wilcoxon Rank Sum Test was used to determine whether there was a statistically significant improvement between the population before the experimental treatments and after. Since the pre and post tests were not paired to the student who completed the surveys, a Wilcoxon Signed Rank Test was not appropriate.

Student confidence in their ability to critique their peers shifted positively over the course of the treatment unit. There was a statistically significant increase in student perception of their ability to offer feedback on work done by their peers that would help their peers reach improved outcomes (Figure 2). The Wilcoxon Rank Sum Test was used to generate a p value of 0.033. This result was also reflected in the student interviews that took place before and after the treatment. Before completing the treatment unit, students were asked “Are you able to be honest with your peers about ways that they can improve?” Student A responded, “No, because I don’t want to hurt anyone’s feelings.” After the treatment, students were asked to respond to the question “How has the unit changed your ability to offer honest feedback to your peers?” Student A said, “I am more honest. There is no point in not being honest because they will just get a bad grade.” This change indicates that the student considered that her critique has value in helping the student she reviewing.

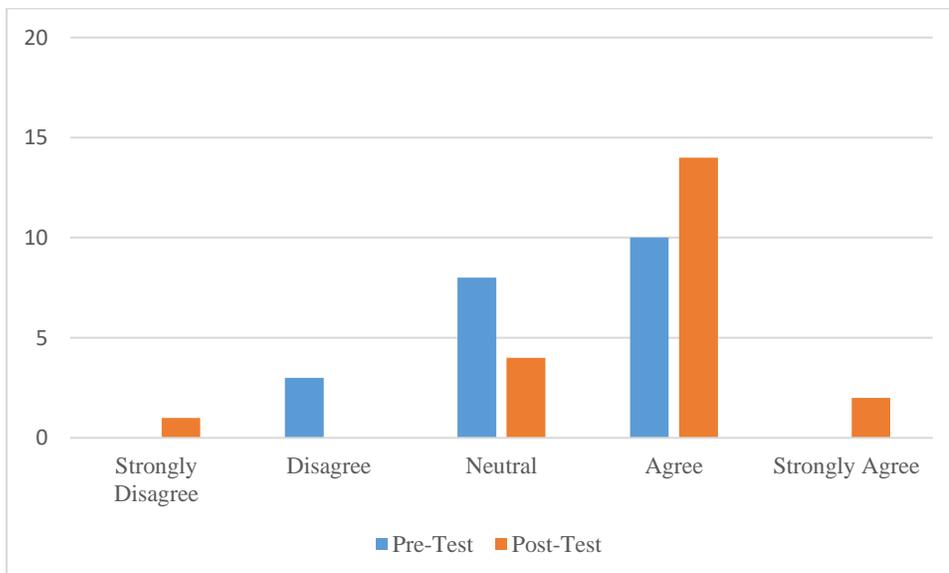


Figure 2. Peer review attitude survey question: “I am able to give other students feedback on their work that will improve their learning outcomes,” ($N=21$).

During group work, students were asked to participate in critiquing how well groups presented an argument on their whiteboards before presenting the information to the class. Students used Post-It notes to ask questions about the information on boards to determine whether students were able to generate useful feedback. The purpose of the board was to communicate how to best treat a Lyme disease epidemic. Students varied in the complexity of the questions that were generated but on each board, there were examples of questions that the students could use to add more depth to their argument. One student asked, “What happens if you get treated with the wrong antibiotic?” Another student asked “Why are there different drugs for adults and children?” In both cases the question leads to more investigation. All students were able to participate in the activity, but some asked questions that were more directed to better supporting an argument with quantitative data. One student asked, “How long does it usually take to spread to your central nervous system?” to indicate to the group that they should add this piece of information in their argument. At minimum, the questions posted on the boards asked for clarification of information and the full participation of the class confirms that the students are confident in their ability to participate in peer review in a low risk setting.

Student views about the importance of feedback on their own work remained consistent before and after the treatment unit. There was not a statistically significant increase in student perception of how important feedback from other students helped to improve their work. The Wilcoxon Rank Sum Test was used to generate a p value of 0.174. The majority of students started the treatment unit with a positive view of the role of feedback from other students. Prior to engaging in the treatment unit 3 students

disagreed or strongly disagreed with the statement “Feedback from other students on my argumentation papers helps me improve the quality of my work.” After the treatment unit, all of the students either strongly agreed, agreed, or were neutral to the same statement as shown in Figure 3.

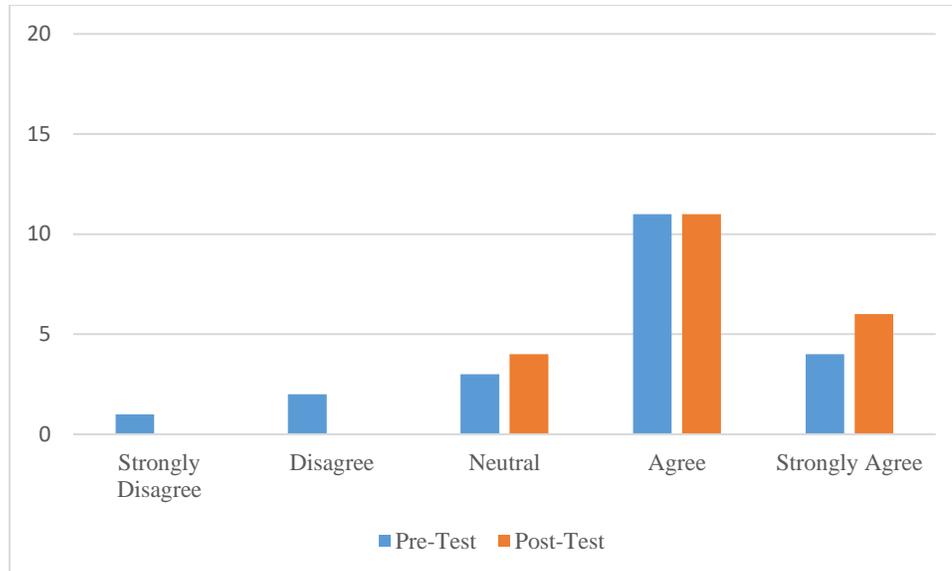


Figure 3. Peer review attitude survey question: “Feedback from other students on my argumentation papers helps me improve the quality of my work,” ($N=21$).

In the Peer Review Attitude Survey and the individual interview, students were asked about the value of the peer review process in making arguments better. There was a statistically significant increase in student perception of the peer review process as a valuable tool to help improve student arguments (Figure 4). The Wilcoxon Rank Sum Test was used to generate a p value of 0.033. This result was also reflected in the student interviews that took place after the treatment. This change in thinking about peer review was indicated by some interviewed students. Some students maintained a positive view of peer review. Before and after completing the treatment unit, students were asked “Do you find peer review valuable? Why or why not?” Prior to the treatment, student A

responded, “No, because it depends on who is editing.” After the treatment, student A said “Yes, because it gives you more ideas.” Students that maintained a positive view of peer review shifted the reason why they valued the process. Before the treatment, student B said, “Feedback is helpful to improve work.” After the treatment, student B said “I value peer review because it is helpful to see the problem from someone else’s perspective.” This change demonstrates a shift in how the student perceived the process.

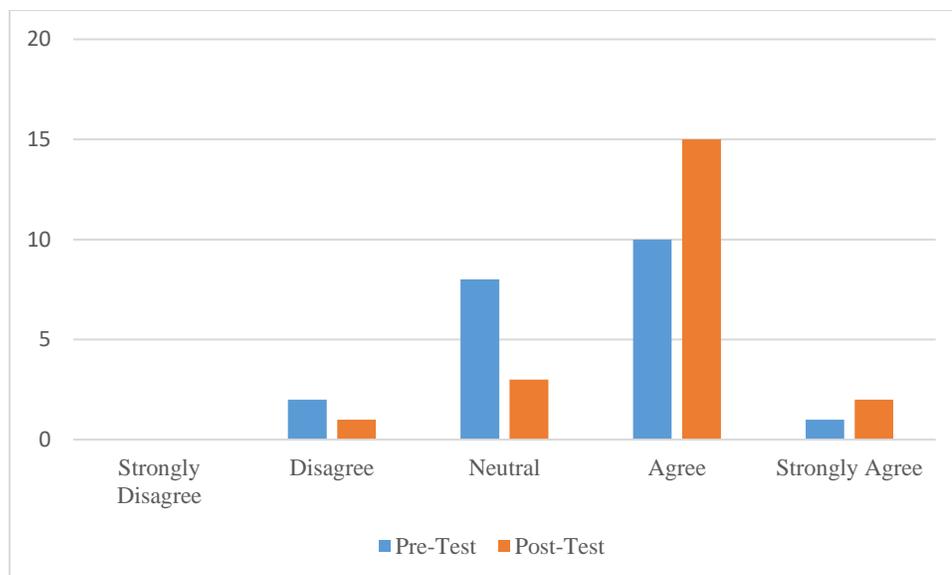


Figure 4. Peer review attitude Survey question: “The peer review process is valuable in improving student arguments,” ($N=21$).

This result was supported when students were completing the self-reflection prompts. When asked, “What makes a good board?” one student wrote, “You need to be able to answer any questions possible that people may ask about your board.” This student identified that preparing for rebuttal is a crucial component of a good argument. In response to the same question, a second student wrote, “If there are any terms or phrases that may cause confusion, make sure it is clearly explained somewhere.” In his response, this student considered the perspective of the viewer in critiquing the clarity of

the argument on the whiteboard. The white boards serve as a way for student groups to share out their current understanding of an argument. Students were able to critique board arguments and then included the perspective of potential peer critics when constructing their responses about what makes a student group white board model good.

Students appeared to remain confident in their ability to participate in the peer review process before and after the treatment unit. There was not a statistically significant change in students' perception of any remaining item on the Peer Review Attitude Survey. Before completing the treatment unit with peer review strategies, the majority of students already agreed with the statement "I am confident in my ability to construct an argument based on scientific data." By the end of the unit, there was very little change in student answers, as shown in Figure 5, and this data generated a p value of 0.363 in the Wilcoxon Rank Sum Test. One student shifted negatively in his confidence to construct an argument. It is possible that as he learned more about what a scientific argument entailed and saw what other students were able to produce, his confidence was impacted negatively. Interviews with students varied in whether the students felt more confident in their building of scientific arguments. Before the treatment, students were asked "How would you describe your ability to make scientific arguments based on evidence?" Student A said, "Medium. If I have evidence that I understand, I will be able to argue with it, but if I don't understand I might sound dumb." When asked "How has the peer review process changed your perspective of building scientific arguments?" after the treatment, Student A said, "Explaining the tick data helped me explain more than I would have." To the same question, Student C said, "I think getting practice helps strengthen

my ability to make scientific arguments. Specifically, it helps by giving more than one perspective before turning in the assignment.” Student C described himself as “pretty ok” at constructing arguments before the treatment. This student identifies additional practice and a second perspective, inherent elements of the peer review process, as valuable in improving responses.

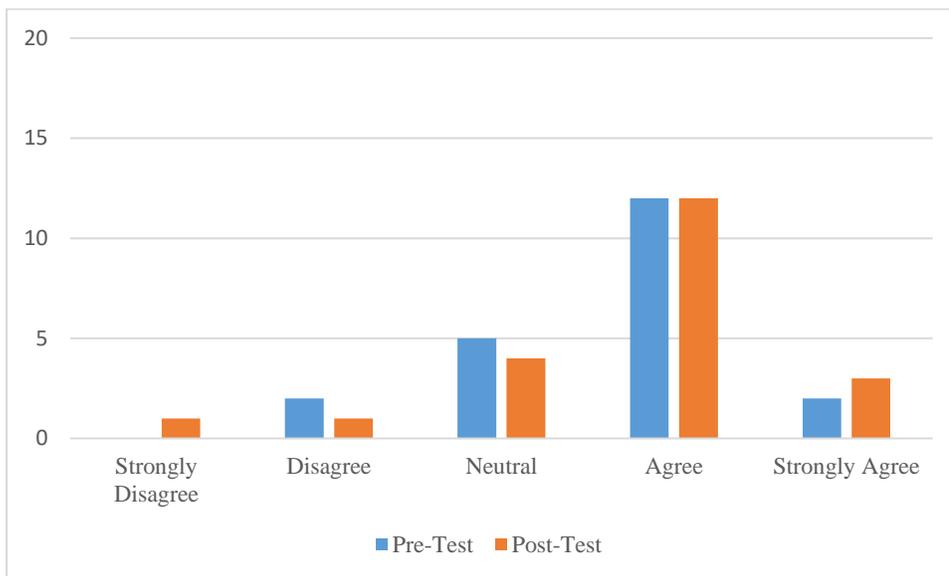


Figure 5. Peer review attitude survey question: “I am confident in my ability to construct an argument based on scientific data,” ($N=21$).

Similarly, the majority of students responded in agreement of the statement “I am confident in my ability to offer constructive feedback to my peers,” as shown in Figure 6. This result occurred before and after the treatment unit indicating little change in student perception of their own ability to provide constructive feedback generating a p value of 0.074. All students that were interviewed before and after the treatment remained consistent in their perception of their ability to offer quality feedback.

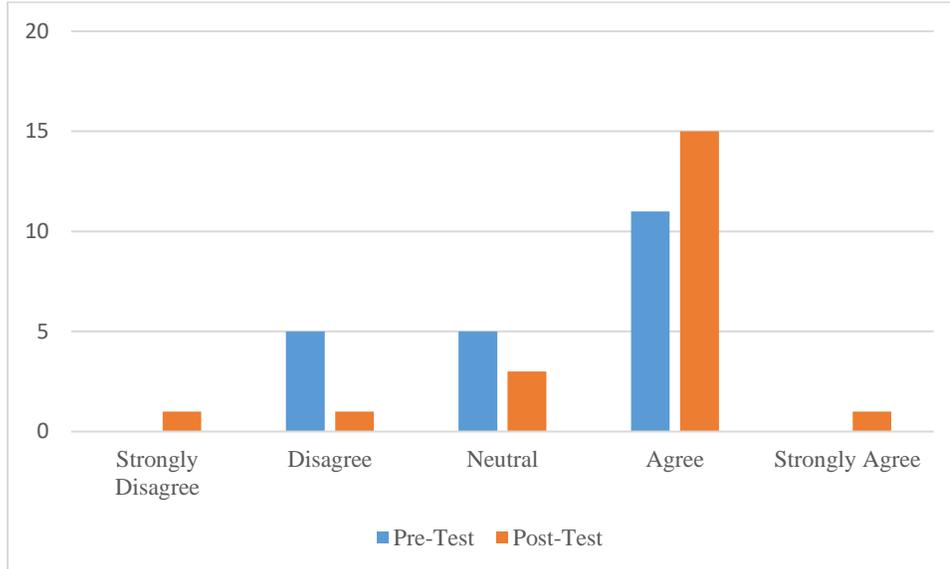


Figure 6. Peer review attitude survey question: “I am confident in my ability to offer constructive feedback to my peers,” ($N=21$).

Student feelings about science did not appear to drastically change during the treatment unit. There was not a statistically significant change in student perception of their own feelings about science nor in their view of whether peer review was used in other classrooms. The number of students that disagreed with the statement “I have negative feelings about science” increased from five to ten, as show in Figure 7, but there was one additional student who strongly agreed with the same statement after completing the treatment unit and these changes generated a p value of 0.166. Similarly, the majority of students agreed with the statement “Peer review is used in other classes” and after the treatment the results were relatively unchanged with a few students more students agreeing and a few less selecting neutral (Figure 8).

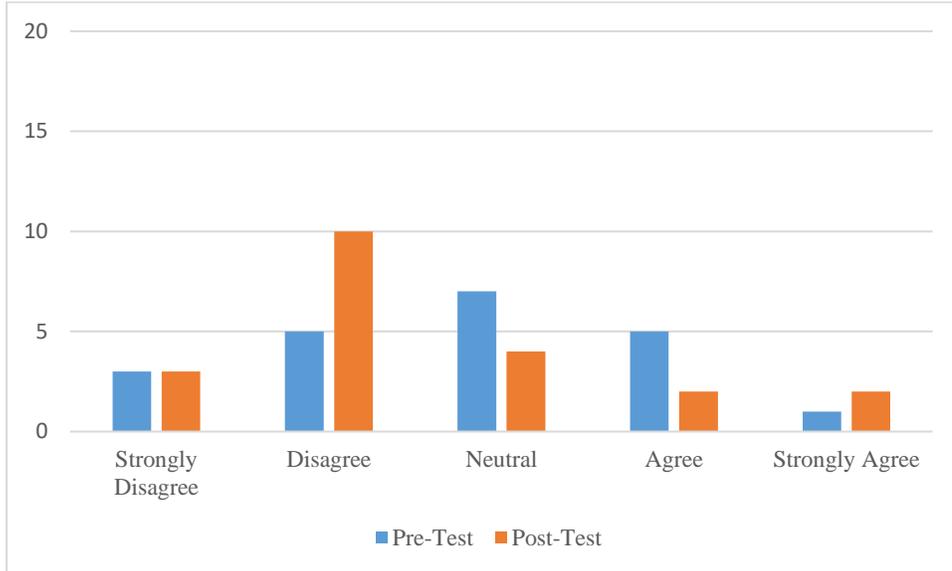


Figure 7. Peer review attitude survey question: “I have negative feelings about science,” (N=21).

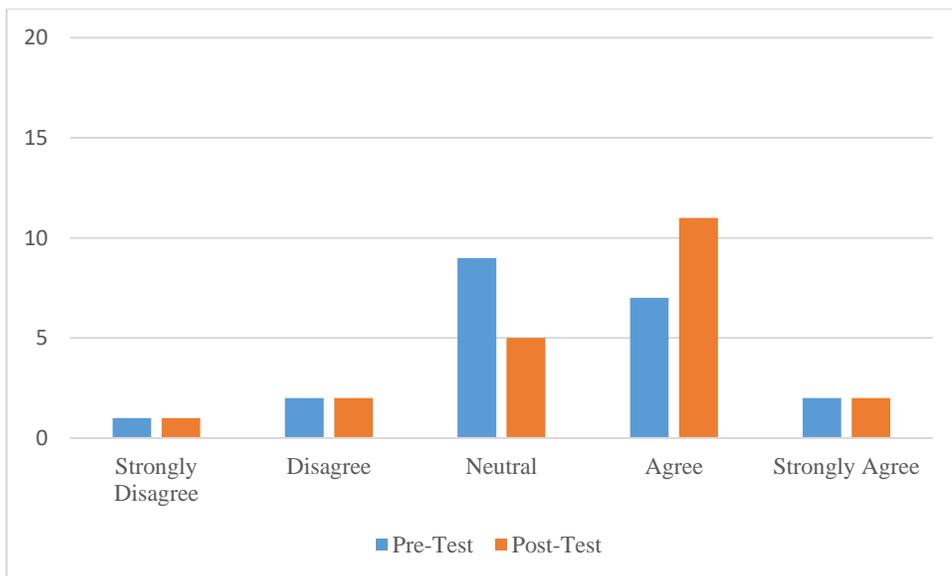


Figure 8. Peer review attitude survey question: “Peer review is used in other classes,” (N=21).

Student feelings about their ability to use evidence to make strong science arguments did not shift drastically over the course of the treatment unit. There was not a statistically significant change in student perception the difficulty of using evidence in argumentation. The majority of students either disagreed with the statement “I find

making arguments based on evidence difficult,” as shown in Figure 9. This generated a p value of 0.417 in the Wilcoxon Rank Sum Test. This was consistent with student interview results. In response to the question “How would you describe your ability to make scientific arguments based on evidence?” student B said, “Pretty strong. The evidence pushes you in a certain direction. Some people just don’t have an interest in it.” Student B rated herself as “pretty strong” at the end of the treatment as well. Students that were interviewed were consistently confident before and after the treatment in their abilities to construct arguments based on evidence.

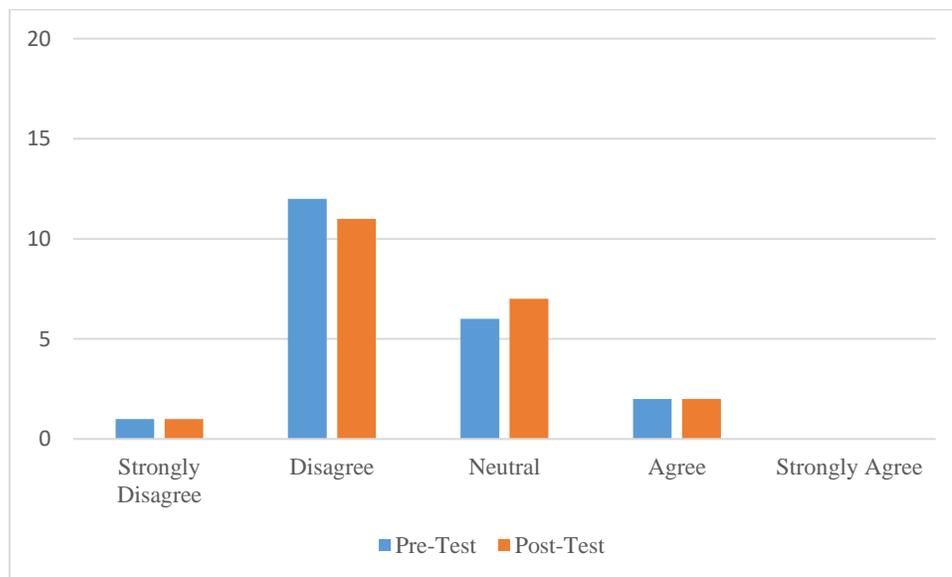


Figure 9. Peer review attitude survey question: “I find making arguments based on evidence difficult,” ($N=21$).

The treatment unit had very little impact on student feelings about where their feedback originates. There were no statistically significant changes in student perception of the comparison between feedback coming from the teacher versus other sources and student perception of their ability to take feedback from others and apply it to their own work. The former generated a p value of 0.330 and the latter 0.337. The majority of the

students disagreed with the statement “The only valuable feedback that is valuable comes from the teacher of my class,” as shown in Figure 10. Student views remained relatively consistent before and after the treatment unit. Similarly, most students agreed with the statement “I can take feedback from others and apply it to my work to make it better,” before and after the treatment unit, as shown in Figure 11. Students that were interviewed were consistently positive before and after the treatment when asked “Do you find it easy to take feedback from your peers and apply it to your work?” Student A had a different experience. She started positive but after the treatment she said, “No because sometimes I do not think I need to change the specific thing the reviewer sees. When they ask good questions that make me think, then it is good.” This response indicates that the student experience can vary drastically based on peer reviewer.

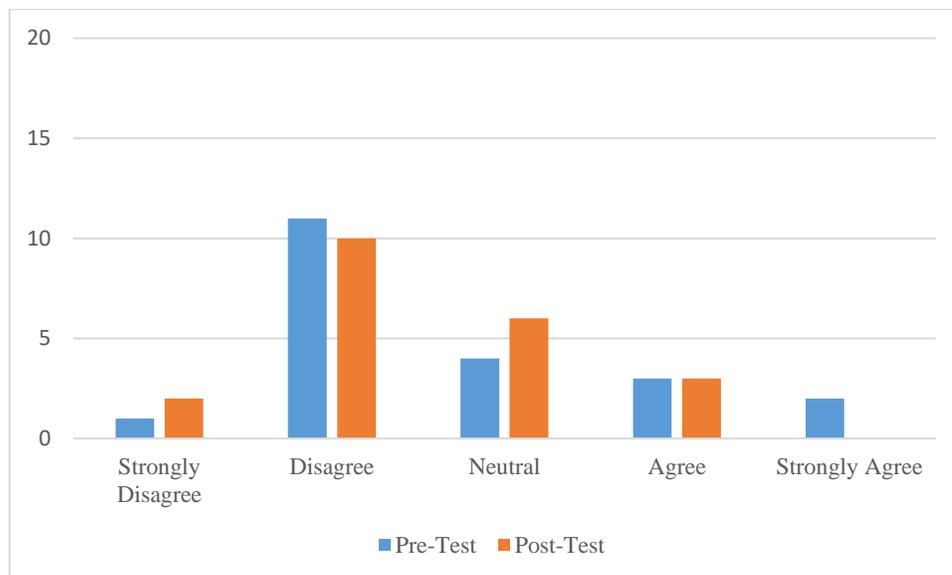


Figure 10. Peer review attitude survey question: “The only feedback that is valuable comes from the teacher of my classroom,” ($N=21$).

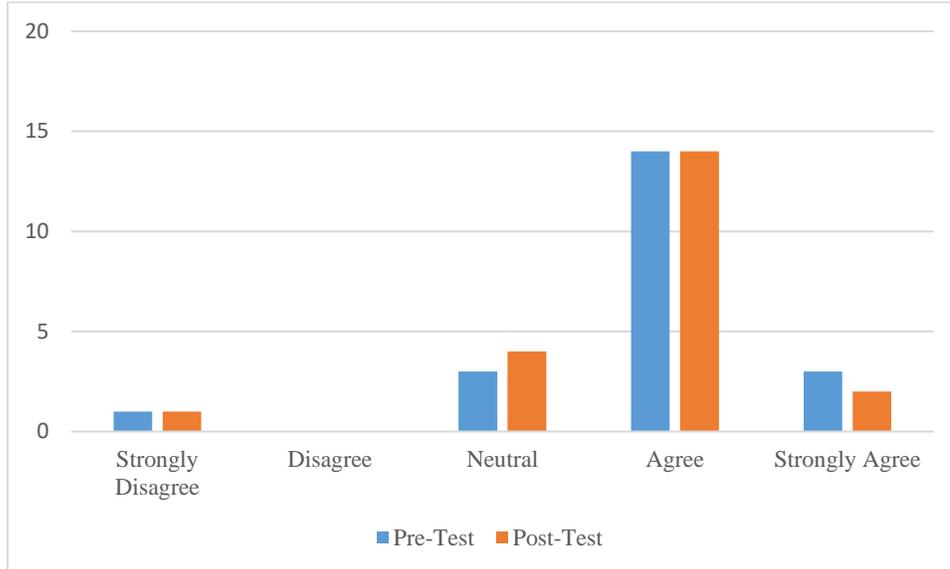


Figure 11. Peer review attitude survey question: “I can take feedback from others and apply it to my work to make it better,” ($N=21$).

The Experimental Design Performance Assessment was given to students before and after the treatment unit to determine changes in how they complete the problem. There was a statistically significant change in the number of students who were able to select a precise dependent variable in the Experimental Design Performance Assessment. Prior to completing the treatment unit, thirteen students were able to select a dependent variable that was more precise than the health of the test organism, which was suggested in the proposed test question in the assessment. After the treatment unit, the number of students who were able to narrow the dependent variable increased to seventeen, as shown in Figure 12. This result generated a p value of 0.021 in a Paired T-Test at $p < .05$ significance. Student success in the task was determined based on whether the student directly stated their dependent variable and whether the dependent variable was some factor other than health of the target organism. A student that selected a more precise dependent variable might indicate that they are observing tick activity levels over a

period of five minutes as a measure of tick health. The feasibility of the dependent variable was not taken into consideration because the students were instructed to complete the problem with unlimited resources.

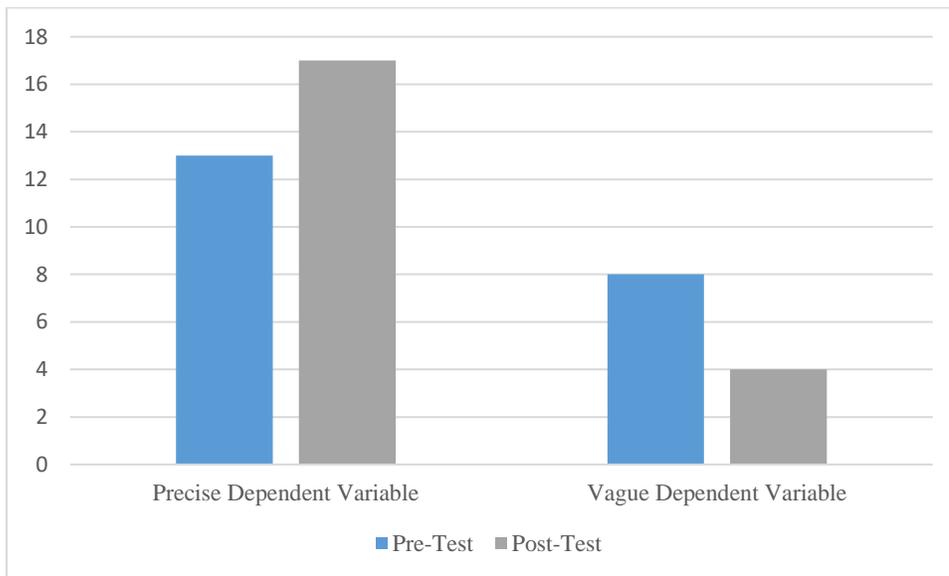


Figure 12. Student choice of dependent variable on experimental design performance assessment, ($N=21$).

Students also selected a variety of dependent variables that represented the health of the target organism in the problem. One student suggested counting the number of ticks that reproduce to indicate tick health and she rationalized this by saying that the ticks would have to be healthy in order to reproduce. This student only incorporated a rational after completing the treatment unit. A second student suggested using weight and activity levels as a measure of tick health. He rationalized his choice of using the two dependent variables together by saying that too little food would lower weight and activity and too much food availability would raise weight but lower activity by making the ticks sluggish. He claimed that only the right amount of food availability would maximize a weight and energy level ratio. Regardless of their choice, students

demonstrated more precision and increased stated rational in their dependent variables after completing the treatment unit. When their selected dependent variable was more precise, the students were also able to extend their explanation for why their chosen dependent variable was valid and were able to make an argument for how their design was fair.

This result was supported by the types of questions students asked when they were assigned to critique boards generated by other student groups using Post-It notes. Many of the questions that students generated were about the data that one might accumulate in order to support a particular course of action when treating a Lyme disease epidemic. When one student group did not include labels on their graph, one student wrote, “What do the numbers in the graph represent?” Furthermore, when a group included a graph without identifying the meaning of the trend, one student asked, “Why is the graph increasing?” This shows that a portion of the student population is considering that identifying data is important when making an argument and that explaining trends can add depth.

When asked to make a driving question board indicating next steps in the case study, students generated questions that demonstrate that they are considering the precise data that needs to be collected in order to support future arguments. Students were asked, “What questions do we still need to answer in order to solve the problem?” This question is referring to the rise in Lyme disease cases and students already identified the organism that makes people sick. In response, one student asked, “Has something in the tick’s environment changed?” Another student asked, “What role do ticks play in the

environment?” Students confirmed that these pieces of information need to be gathered in order to make our evidence sufficient in answering the question, “What do we do next to slow or stop the spread of Lyme disease?” These student responses during group work demonstrated a possible connection to the growth indicated by the Experimental Design Performance Assessment pre- and post-test results.

INTERPRETATION AND CONCLUSIONS

Peer review and self-reflection are essential skills in science fields but their role is often diminished in science curricula. If students seek out professions in science and engineering, they will be faced with tasks where they will be expected to give and receive precise feedback. As educators of future scientists and a public that increasingly needs to digest scientific content, we can do better to prepare students for an environment where they will have to seek out best practice by collaborating with peers and where their conclusions will be openly challenged. Unfortunately, much of the practice provided in a traditional science classroom does not match the reality of the practices used by professionals in science careers. Adding peer review and self-reflection strategies into my curriculum may serve as a solution to fill this void in that it has the potential to assist students as they calibrate their skepticism of designed experiments. Furthermore, practicing peer review helps students consider the perspective of potential critics and this can lead to more precise selection of dependent variables when designing their own investigations. Lastly, engaging in frequent peer review can strengthen student confidence in the role it plays in strengthening scientific arguments and in their own ability to provide meaningful feedback.

The increase in student skepticism of the conclusion drawn by the fictional scientist was particularly encouraging in that students were more critical of what they were reading after critiquing each other on designed experiments. Science thrives on peer critique to analyze collected data from possible angles to get as close to the truth of what occurs in a phenomenon as possible. If there is a breakdown in the level of skepticism that future scientists practice then predictions based on data will suffer in their precision and accuracy. Increasing student skepticism is vital to their development in constructing arguments based on the claim evidence reasoning framework, where there is a higher than typical standard for evidence selection. The results of this portion of my investigation match the findings in previous studies. In helping the students see themselves as practitioners that can apply information that they learn from experimentation, we can help support students as they begin to use gathered evidence to make more accurate predictions about a related phenomenon.

At the end of the treatment, there was a statistically significant increase in the number of students that were self-critical enough to select a more precise dependent variable when given a fictional research question. They took the implied dependent variable in the given question, “tick health,” and selected a variable that could represent a measure of tick health. When students did this, they were able to write a more substantial defense of their experimental design by naming more variables that they controlled to keep the experiment fair and a more detailed defense of the dependent variable that they selected. As freshmen in high school, these students are just beginning their journey into feeling as though they are contributing members of the scientific community. A crucial

step in this journey is developing an internal dialogue that helps maintain a high level of self-evaluation which later helps the student when defending the data that is selected as evidence of a particular claim made. This is an independent journey and there was no direct instruction on how to avoid selecting an imprecise dependent variable and implying a defense. The students demonstrated that they were more critical of “health” as an imprecise variable and they reported that they felt they could give feedback that would help other students improve experiments.

Student confidence in the peer review process increased after engaging in the peer review strategies on a daily basis. This finding mirrors the results of similar investigations at different academic levels. Despite previous research, I expected a certain level of fatigue in using peer review strategies after ten days of coursework but the students either maintained a positive perception of peer review or their perceptions increased. Students saw the process as more valuable in making arguments in science better after completing the treatment unit. They also reported a higher level of confidence in their ability to offer feedback that could help make peer outcomes better.

Based on the information generated in this investigation, peer review will serve a greater role in my classroom. Specifically, the strategies that I used in the treatment unit, along with the case study format, appear to be especially valuable in helping students understand that peer critique is vital to maintaining the academic standard of the scientific community. Furthermore, practicing as a peer editor and receiving frequent feedback from peers will help my future students make more precise decisions as they practice designing investigations of their own and as they draw conclusions from results

from other studies. Peer review will be built into my future curriculum so that my future students can benefit from multiple perspectives when learning how to construct arguments based on evidence.

VALUE

The strategies implemented in my classroom generated encouraging results in that the students demonstrated a higher level of skepticism about experimental design and were more likely to select evidence that was measurable and defensible. These skills are crucial if the students are to make progress in developing arguments based on the claim evidence reasoning framework. Additionally, student surveys demonstrated that the students perceived the peer review processes provided in our class as valuable in helping to achieve higher goals and in building their confidence in their ability to provide useful feedback. These results indicate that frequent use of peer review strategies are a viable strategy for helping my students take their understanding of how science functions to the next level. In practicing as the reviewer and the reviewee, the students will obtain skills that are vital for anyone living in a modern society where science content is frequently deployed as evidence for action.

Early in my career, I chose to adopt the claim evidence reasoning framework in my classroom, but I had limited guidance on how to help students get better at using evidence in their arguments. I wanted to teach the students how to arm themselves with data before attempting to construct an explanation but I found myself repeating the same lines to the students with little success. Learning how to argue and defend claims effectively was not a skill that could be captured completely in lecture; students needed to

experience the success of a well-supported argument compared to the failure of vague claims supported by inappropriate or insufficient evidence. I sought to find a way to expose students to the experience in a low risk setting and peer review was a promising candidate. Specifically, students needed a format where they could calibrate their critical eye when given a set of data without being heavily penalized for their present understanding. Furthermore, I wanted my students to improve in their ability to select appropriate evidence to gather before carrying out a designed experiment. This also requires a critical eye and multiple perspectives. Being able to identify the appropriateness of data is a crucial step before students can take valid conclusions and generalize to include related phenomena. Growth in developing strong scientific arguments is best experienced with their peers when the stakes are low.

Peer review is an excellent way for students to contemplate other ways of approaching a scientific argument and it is a dynamic way to help students as they are developing an analytical mind. Considering other perspectives is crucial when it comes to constructing and defending strong arguments based on the claim evidence reasoning framework. The best arguments are made as the writer is predicting what others will say about the argument once it is out in the open. Students who practice peer review in the science classroom are more likely to consider what others might say in dissent. It appears that the peer review strategies that were used in the treatment unit were effective in helping students understand this benefit of the peer review process. Students identified the importance of multiple viewpoints in the student interviews, surveys, and while they performed activities. Students approached their duty to critique their peers seriously and

saw the exchange of ideas as an opportunity to practice developing an argumentation structure. Establishing that peer review is a valuable strategy to maintain the integrity of science as an endeavor helps humans reach beyond their individual biological limits.

In completing this action research, I have renewed confidence in using low risk peer review in all of my classes. Over the course of my career, I continually tried to find the next big technique to add to my classroom and the action research process forced me to reconsider a technique I half-heartedly tried in the past. Before starting, I was looking for an excuse to use peer review in a more purposeful way. I identified that my students required more practice reflecting on their experience and during the peer review process, students continually express themselves outward. After seeing the results of my action research, I confirmed that peer and self-review can serve a role in my classroom without taking a tremendous amount of time away from other practice that might be meaningful to subsets of students. I was surprised at how fluid peer review can become in the classroom with practice on my end.

Successful use of the claim evidence reasoning framework was difficult to quantify. I chose to focus on whether students were taking steps toward supporting their arguments with appropriate and sufficient evidence, but in the future I will not be satisfied until students are making arguments that reason through why they think their evidence is sound in relation to the claim. Furthermore, I was surprised at how much questioning was required before students started to increase their criticism of scientific practices. I will continue to practice identifying student needs with regard to using good scientific argumentation. Guiding students as they become more confident in their

criticism is a rigorous process and I was caught off guard how much revisiting needs to take place when transitioning to new scenarios. Regardless, I am committed to the process as it is so important to have a citizenry that is capable of reading and understanding information that the scientific community makes public.

Before beginning the action research process, I had several concerns about using peer-review more openly in my classroom. I was worried that adding more peer review strategies would slow instruction to a crawl and that students would perceive the tasks as tedious. In the end, it was clear that students saw the process as valuable and the time taken for peer review did not have a large impact on the flow of the classroom. On the contrary, whiteboard presentations to the class went faster because students addressed concerns from their classmates before their presentation and audience members already had questions in mind for each group presentation. Asking helpful questions takes practice and if this becomes part of my typical routine in the classroom, then the time spent on peer review practice will likely be saved in the long run as students make more distinct content gains in a shorter period of time.

At the start of my action research, I had similar concerns regarding argumentation in the classroom. In taking data directly on gains in using evidence in argumentation, I learned that it can be frustrating to get students to use data in their argumentation practices. At the end of my units that did not contain additional peer review, students were struggling to identify measurable dependent variables and flaws in fictional experiments. I was shocked by this outcome because students get direct practice in both of these skills in their first experimental design unit. I was concerned about retention of

these newly acquired skills and peer review seemed like a logical choice to help reinforce practice that students receive indirectly in later units. It turns out that peers can be an valuable resource in helping each other gain experience in essential science practices. Furthermore, I was concerned initially that students needed more practice engaging in polite discourse and in the future I will keep better track of how the students are communicating with each other. I expect that students who use arguments that are data driven tend to engage in discourse more politely without conceding ground but I will not know for sure until the strategies are implemented in my classroom universally.

After consideration of the data generated in my action research, I will be expanding the role that peer review plays in my classrooms. Peers can be valuable resources as students test their current academic limits but only if students have practice trading feedback with classmates. Written argumentation is a particular area of concern in my classroom. As many students identified in my interviews, it is valuable to consider multiple perspectives because there are some moments when perception is all that matters to the public. I will likely include more peer review and self-reflection across the entire semester, but I will avoid concentrating all of the experiences into a small time period. I am still afraid of the students becoming fatigued with the peer review process. Furthermore, I will spend more time having students practice arguing about the validity and sufficiency of data. This provides a perfect environment for students to get practice responding to challenges from peers that further refines scientific understanding of a particular phenomenon.

REFERENCES CITED

- Blikstein, P., Fuhrmann, T., & Salehi, S. (2016). Using the Bifocal Modeling Framework to Resolve “Discrepant Events” Between Physical Experiments and Virtual Models in Biology. *Journal of Science Education and Technology*, 25(4), 513-526.
- Fortus, D., Schwartz, Y., & Rosenfeld, S. (2015). High School Students’ Meta-Modeling Knowledge. *Research in Science Education*, 46(6), 787-810.
- Geithner, C. A. & Alexandria, A. N. (2015). Doing Peer Review and Receiving Feedback: Impact on Scientific Literacy and Writing Skills. *Advances in Physiology Education*, 40(1), 38-46.
- Gjalt, P. T., Bulte, A. M. W., & Pilot, A. (2016). An Activity-Based Instructional Framework for Transforming Authentic Modeling Practices into Meaningful Contexts for Learning in Science Education. *Science Education*, 100(6), 1092-1123.
- Glaser, R. E. (2014). Design and Assessment of an Assignment-Based Curriculum to Teach Scientific Writing and Scientific Peer Review. *Journal of Learning Design*, 7(2), 85-104
- Kelly, L. (2015). Effectiveness of Guided Peer Review of Student Essays in a Large Undergraduate Biology Course. *International Journal of Teaching and Learning in Higher Education*, 27(1), 56-68.
- McNeill, K. L., & Pimentel, D. S. (2009). Scientific Discourse in Three Urban Classrooms: The Role of the Teacher in Engaging High School Students in Argumentation. *Science Education*, 94(2), 203.
- McNeill, K. L. & Martin, D. M. (2011). Claims, Evidence, and Reasoning: Demystifying Data During a Unit on Simple Machines. *Science and Children*, 48(8), 52.
- Schen, M. (2013). A Comparison of Biology Majors’ Written Arguments Across the Curriculum. *Journal of Biological Education*, 47(4), 224-231.
- Shemwell, J. T., Gwarjanski, K. R., Capps, D. K., Avargil, S., & Meyer, J. L. (2015). Supporting Teachers to Attend to Generalisation in Science Classroom Argumentation. *International Journal of Science Education*, 37(4), 599-628.
- Walker, J. P., & Sampson, V. (2013). Argument-Driven Inquiry: Using the Laboratory to Improve Undergraduates’ Science Writing Skills through Meaningful Science Writing, Peer-Review, and Revision. *Journal of Chemical Education*, 90(10), 1269-1274.

APPENDICES

APPENDIX A
EXPERIMENTAL DESIGN PERFORMANCE ASSESSMENT

Name _____

1. Design an experiment for the following question. **Use pictures, graphs, and words.**
Does food availability impact tick health?

Please include the following. . .

- i. A specific independent variable.
- ii. A control group and an explanation for why it is the control group.
- iii. A measurable dependent variable.
- iv. A fair and controlled experimental design.

APPENDIX B
RUBRIC FOR EXPERIMENTAL DESIGN PERFORMANCE ASSESSMENT

Rubric

This assessment includes an unscientific question so that students must come up with ideas for how to appropriately and sufficiently measure a dependent variable. The students must then make an argument defending their experimental design.

A specific independent variable.

2-Independent variable is a change in food availability between groups and includes many subjects

1-Independent variable is a change in food availability but does not consider sample size

0-No clear change in independent variable

A control group and an explanation for why it is the control group.

2-Control group is identified and serves as a basis of comparison to eliminate other possible factors

1-Control group identified but purpose not clearly explained

0-No evidence of control group

A measurable dependent variable.

2-Dependent variable is an appropriate measure of tick health

1-Dependent variable is loosely connected to tick health

0-No evidence of a dependent variable connected to tick health

A fair and controlled experimental design.

2-All variables other than enclosure size are kept controlled and several are mentioned

1-Many variables are kept controlled but changes may result in another variable also changing

0-Changes to more than one independent variable.

Number 2

2-Students have developed a measure of tick health and make an argument for why it is more appropriate than other measures

1-Students have developed a measure of tick health but do not include an argument for why it is more appropriate than other measures

0-Students are unable to develop one dependent variable that measures

Number 3

2-Students identify several controlled variables that eliminate other possibilities that may contribute to differences in health

1-Students identify several controlled variables but they do not eliminate possibilities that may contribute to differences in health

0-Students are unable to identify several variables that are controlled in their experiment.

APPENDIX C

EXAMPLE OF AN EXPERIMENTAL DESIGN EVALUATION EXIT SLIP

Name _____

A local entomologist just discovered a new tick species. This tick species appears to seek out mammalian hosts, where they get a blood meal to sustain them as they continue their life cycle. The entomologist sets up an experiment to determine how the new tick species locates its host. The experimental setup contains a tube with a clear boxes at each end. In order to determine whether the ticks use heat to detect their host, the scientist shines a heat lamp on one of the boxes and puts 10 ticks (5 females and 5 males) directly in the center of the tube. After allowing the ticks to move around for two minutes, the scientist records the distance the ticks traveled from their starting position in the center of the tube. The data is recorded in the following table.

Tick #	Distance from starting position (cm)	Direction
1	0	N/A
2	5	Toward heat lamp
3	18	Away from heat lamp
4	7	Toward heat lamp
5	0	N/A
6	3	Toward heat lamp
7	8	Toward heat lamp
8	7	Toward heat lamp
9	2	Away from heat lamp
10	10	Toward heat lamp

The scientist concluded that ticks use heat to detect their hosts based on the data she collected.

Do you agree with the conclusion that the scientist made? Defend your response.

APPENDIX D
PEER REVIEW LIKERT SURVEY

Name _____

Student Survey

For each of the following items, indicate to what degree you agree or disagree with the statement.

1. I am able to give other students feedback on their work that will improve their learning outcomes.

Strongly Disagree Disagree Neutral Agree Strongly Agree

2. Feedback from other students on my argumentation papers helps me to improve the quality of my work.

Strongly Disagree Disagree Neutral Agree Strongly Agree

3. The peer review process is valuable in improving student arguments.

Strongly Disagree Disagree Neutral Agree Strongly Agree

4. I am confident in my ability to construct an argument based on scientific data.

Strongly Disagree Disagree Neutral Agree Strongly Agree

5. I am confident in my ability to offer constructive feedback to my peers.

Strongly Disagree Disagree Neutral Agree Strongly Agree

6. I have negative feelings about science.

Strongly Disagree Disagree Neutral Agree Strongly Agree

7. Peer review is used in other classes.

Strongly Disagree Disagree Neutral Agree Strongly Agree

8. I find making arguments based on evidence difficult.

Strongly Disagree Disagree Neutral Agree Strongly Agree

9. The only feedback that is valuable comes from the teacher of my class.

Strongly Disagree Disagree Neutral Agree Strongly Agree

10. I can take feedback from others and apply it to my work to make it better.

Strongly Disagree Disagree Neutral Agree Strongly Agree

APPENDIX E
STUDENT INTERVIEW QUESTIONS

Student Interview Questions

Pre-assessment

1. Have you ever been in a class where peer review was commonly used to make your work better? If so, can you describe how it is used?
2. Do you find peer review valuable? Why or why not?
3. Do you think you can give other students feedback that will help them to get better?
4. Are you able to be honest with your peers about ways that they can improve? Why or why not?
5. Do you find it easy to take feedback from other students and apply it to your work? Why or why not?
6. How would you describe your ability to do science?
7. How would you describe your ability to make scientific arguments based on evidence?

Post-assessment

1. How did you like how peer review was used during this class?
2. Do you find peer review valuable? Why or why not?
3. How has this unit changed your views about the value of peer review?
4. Do you find it easy to take feedback from your peers and apply it to your work? Why or why not?
5. How has the unit changed your ability to offer honest feedback to your peers?
6. How has the peer review process changed your perspective of building scientific arguments?
7. Has the peer review process changed your thoughts about your science education? Why or why not?