CLAIM, EVIDENCE AND REASONING IN THE SCIENCE CLASSROOM

by

Amy Leigh Steele

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 2021

DEDICATION

This work is dedicated to my family and friends. My family has been supportive in every way possible, particularly my husband Chris Steele. He has filled in many gaps and carried unexpected loads. My daughter Olivia has been a cheerleader and a helper in everyday life, particularly during the exceptional spring, summer and fall of 2020. My daughter Nora has been a patient and helpful person during this process. My parents Ann and Larry Heard valued education and modeled life-long learning for their children. Finally, I would like to dedicate this project to the memory of my son, John Patrick "Jack" Steele, my fellow Bobcat and science nerd who was delighted to have his mom in school with him and whom we miss terribly every day.

ACKNOWLEDGEMENTS

Cornerstone Christian School has been my teaching home for the past decade; it is an exceptional small school where students are supported and challenged as well as a place where teachers are both expected to innovate and allowed to meet the needs of their students in terms of pedagogical practice. I could not have done this without my smart and delightful middle school students who are learning to think like scientists more every day. I think they know how highly I regard them and their potential to make a difference in this world. Our administration, school board and parent community have been supportive of this effort in extraordinary ways. I cannot thank them enough for their kindness and accommodation. Every teacher should be so lucky.

To my friends and peers Lori Stanton and Mare Sullivan, thank you for innumerable conversations, stupendous suggestions and professional upbraiding at times. Colleen Boye did a wonderful job illustrating Appendix B; she is a talented author and illustrator who interprets her mother's work and by extension, her mother's friends in their endeavors with generosity and creativity.

Lastly, John Graves has been a blessing as an advisor. He is a gifted leader to teachers, who are not always the easiest to instruct. His contributions and encouragement to this project kept it afloat through some turbulent times. His work embodies the goal of the MSSE program to make every teacher a better teacher.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

ABSTRACT

This Capstone Project's purpose was to examine the effects of an educational approach called Process-Oriented Guided Inquiry Learning (POGIL) on student abilities to create, identify, and understand different aspects of scientific argumentation. The research followed two middle school classes through the process of treatment, and inquiry labs where they engaged in argumentation in response to a guiding question. Students in grades six and seven investigated the differences among claim, evidence and reasoning in scientific argumentation. The research indicted low gains in identifying examples of argumentation components and creating written argumentation for both classes.

CHAPTER ONE

INTRODUCTION AND BACKGROUND

Context of Study

In my first few years as a science teacher, students were taught the scientific method as the standard for inquiry in science. This idea informed science fair rubrics for grading and judging, students were expected to know its components by heart, and they were taught how to use it as a means to answer questions. Over time, I observed that students often had excellent questions to explore that didn't fit neatly into the scientific method scaffold for investigation. Together we often bent ideas around that scaffold to conduct inquiry. Students quickly saw that not every investigation worked well with the scientific method. It became one of many possible ways to understand and approach an inquiry problem.

These experiences led me to research activities and instructional opportunities which could lead to greater growth in scientific thinking and argumentation. A broader set of values for communication of ideas in science emerged along the lines of the Claim, Evidence and Reasoning (CER) framework. Unsurprisingly, this change came with its own set of challenges in implementing those new values successfully.

Middle school students could often identify a valid and solid claim based on pieces of evidence and patterns they observed in data and models. This was a welcome and more intuitive way to communicate science findings. However, most students struggled to present the reasoning behind the evidence. Their efforts were usually a restatement or mixture of the claim and the evidence, not the reasoning behind it. I incorporated new strategies in the form of argument driven inquiry labs, but found the same problems cropping up. I began to wonder what I could do as a teacher to offer students a chance to gain proficiency in this area of reasoning.

While doing background research, I discovered that many other middle school teachers had encountered similar issues. After leading several units of argument-driven inquiry, I was still seeing my students still struggle to separate evidence from reasoning as well as not be able to present a well-developed argument for their claims. I wondered how to model best practices using techniques that weren't too challenging. Students needed some type of scaffolding to help them grasp the difference between the evidence supporting a claim and the reasoning that explains the link between the claim and the evidence itself. I decided to use a more structured treatment approach via a tool which had high engagement and interest with my students already.

Process-Oriented Guided Inquiry Learning (POGIL) uses teacher-facilitated, small group activities to introduce students to new ideas and vocabulary in science and mathematics through structured models and questions. Most POGILs use multiple models with increasing complexity to compare and contrast concepts. I often used POGIL in my classroom to explore new fundamental science ideas and concepts as well as model abstract ideas. I wondered if I could develop and use a specific POGIL alongside my argument driven inquiry units to introduce and break down components of claim, evidence and reasoning.

This seemed very practical on many levels. I have used POGIL for several years with excellent results. It had a high familiarity level as it is a tool my students use early and often in my classes, so bias towards a new instrument would be minimal. Most of my students when surveyed informally about things they have enjoyed each year mentioned POGIL as being a favorite learning activity. The cooperative and small group nature of POGIL allowed me as an instructor to check in with each student or group frequently at specific checkpoints. The three CER elements by nature seemed to lend themselves to POGIL traits such as relatable modeling, comparison and contrast, and identification of good practice among varied examples, like a misconception probe. Thus, the idea for this study came from a desire to see students improve scientific argumentation skills by engaging in guided inquiry and a modified argument driven inquiry. Both Process-Oriented Guided Inquiry Learning (POGIL) and Argument Driven Inquiry (ADI) have their roots in inquiry but use different techniques to engage students.

I was the sole science teacher for my students throughout middle school, I wanted to make sure that I was doing the very best I can for them in meeting their academic needs in preparation for high school and beyond. I hoped to incorporate POGIL as the missing piece in allowing students to put the whole puzzle together or improve CER skills.

School and Student Demographics

Our school is a small, private school with a diverse student population. Cornerstone Christian School is in Lacey, Washington. In 2020-2021 our school served 142 students in grades K-8. Our ethnicity was Caucasian at 65%, 13% African American, 12% Latino, 9% Asian, 4% Native American, 2% identifying as multiracial, and 2%

Native Hawaiian. Some students identified as two categories, resulting in the percentages being more than 100%. An unusual, consistent aspect of our school was that roughly ten percent of our students were adopted (T. Davis, personal communication, October 2020).

Cornerstone is a private, non-profit, ecumenically diverse Christian school. It grew from a homeschool cooperative in the early 1990s and became an accredited school in 1997 with a single class of first grade students. In 2020-2021, twenty-five percent of our students received financial assistance to attend, and several families qualified for free or reduced school lunches. Two of our families lived in Pierce County just to the north of our school and the rest resided in Thurston County. Our students were placed in single class grades with a cap of 18 students per class. Middle school students had science with me for three years, allowing continuity and community in adolescence when student need for connection is high (Davis, 2020).

The middle school population represented five ethnic groups along with three students who identify as multiracial. The group was also neurologically diverse; two students identified as being on the autism spectrum and six received accommodations both formally and informally in support of diagnosed learning disabilities (ADD/ADHD, dyslexia, and dysgraphia). Four of those six students had diagnosed anxiety or other mood disorders as well. Seven students in middle school were adopted, five internationally. There were two sets of twins as well as half-siblings who are only a few months separate in age and in the same class. Over a third of our middle school students were new to our school this year, possibly a result of COVID school shutdowns and parent concerns over school closures. Although the school average on the Iowa Test of

Basic Skills is in the high 80s in terms of percentiles, the middle school classes were in the 75th percentile in comparison to the school as a whole entity (Riverside, 2010).

Focus Statement

My endeavors to improve critical scientific thinking over the past few years brought me to this focus for my action research project, How will using POGIL affect student proficiency in making their own claim, evidence and reasoning statements in response to a guiding question?

This generated the following sub-questions:

1. Will student abilities to discriminate between claim, evidence and reasoning statements encountered elsewhere improve as a result of POGIL?
2. Will student abilities to differentiate between evidence supporting a claim and evidence that refutes or has no correlation with a claim improve as a result of POGIL?

CHAPTER TWO

CONCEPTUAL FRAMEWORK

POGIL

Process Oriented Guided Inquiry Learning (POGIL) is a structured cooperative learning approach to introducing concepts and vocabulary within a small group using defined roles. One of the early implementations of POGIL was in large group undergraduate chemistry classes at Franklin and Marshall College with high drop rates and low pass rates for students. Class format was traditional lecture with recitation hours led by graduate teaching assistants. Chemistry professors Richard Moog and James Sherman implemented this cooperative learning technique to determine if it could boost student success and pass rates. They created two class formats: one using POGIL in small groups with graduate teaching assistants acting as facilitators and the other a traditional lecture with recitation hour, then ran the two formats simultaneously for four years. What they discovered was overall class pass and non-attrition rates were significantly higher in the POGIL section than in the traditional lecture/recitation model (Moog, 2006).

Moog was influenced to move from a traditional lecture-centered class to a learner-centered approach after observing students struggle to master concepts that he felt they were capable of conquering. His changes were based on learning research, but he also wanted to emphasize process skills which would enable his students to be successful in future situations regardless of pedagogy. In researching how people learn, inquiry came out as a clear best practice. Moog found that inquiry worked best in small groups where students can learn to share their ideas as well as listen to the perspectives of others.

Moog wanted to work within a constrained set of concepts while providing a framework for his students. This led to guided inquiry based on the Karplus learning cycle: exploration, concept development and application. POGIL activities are designed to be used in a content-specific setting and guiding questions are chosen carefully to set students on a course towards mastery. As POGIL has expanded to other levels and branches of science and mathematics, guided inquiry has held up as a format across disciplines. POGIL activities are used from middle school science classrooms to post-secondary institutions (Moog, 2019).

Cooperative learning methods have roots in the idea that obtaining and use of knowledge in practice has strong correlation to social interactions (Christensen et al., 1991). Additionally, students who ask questions, discuss, engage in dialogue, and interact with new concepts in a social setting tend to retain the ideas (Bransford et al., 1999). Background research on teams of learners has some salient points on inclusivity. Cooperative learning in math and science tends to be more supportive for second language learners, female students, and other groups that are not widely represented in science career paths. Reducing isolation and making environments less competitive by promoting collaboration gives everyone in a team a voice to seek clarity with little risk (Bransford et al., 1999).

Most POGIL activities have three main phases echoing a modified learning cycle. It starts with exploration using models to build a construct. Questions about models are structured to guide students to a correct perspective on a core idea while encouraging growth in science and engineering practices. The questions and models begin with everyday examples that students encounter, a story about universal experiences or simple graphics organized for exploration. The second phase moves towards concept ideation and pattern recognition. The final section requires students to apply ideas and patterns to more diverse or complex problems. Students are often required to generate a further response either as a group or individual that reflects on areas where questions remain about a concept. This is often a valuable tool for the instructor as it gives formative feedback (Kussmaul, 2019).

A POGIL is not a complete 5E learning cycle. It is meant to develop science process skills in tandem with exploration of a new concept and give students scaffolded practice in those skills while setting an accurate knowledge foundation. Breaking down an idea into component parts allows students to manipulate those parts in a small collaborative setting so that they may have a stronger connection to the learning. This gives students a solid start to extending and applying their concept base through being able to understand what a concept means and also identify what it is not (Moog, 2006).

Claim, Evidence and Reasoning

Claim, Evidence and Reasoning (CER) has its origins in the manner that scientists work, present, argue and rebut their ideas in a community of peers. Many student populations over the years have been introduced to the scientific method of question, hypothesis, research, experiment, collect data, and draw conclusion as the gold standard for scientific argument. However, the nature of science is much more of a web or a progression than a lockstep method. CER seeks to replicate argumentation in order to create opportunities for students to experience thinking as a scientist in community (McNeill & Martin, 2011).

Cognitive research indicates that students learn science best when given opportunities to work within social structures and a framework of argumentation. CER works best when it becomes part of common vocabulary used by teachers to discuss ideas with students. Students frequently make claims in response to queries about the underlying processes or causes of phenomenon encountered in science investigations. When instructors point out claims that students are making, they begin to see claims as answers to science questions, or inquiry (McNeill & Martin, 2011).

Students in middle school are beginning to interact with data in many different ways, from graphs and tables to visual organizers and illustrative figures. They may use models or photographs of life cycles in organisms. They are able to understand that data in many forms can act as support evidence for a claim, even if they don't always understand how to optimally organize or use what they are seeing (McNeill & Martin, 2011).

The most difficult part of CER for middle school and older elementary students is the reasoning aspect. In order to have sound reasoning, you have to be able to apply known scientific processes, theories or laws to explain how your data supports your claim. Identifying which science ideas are at work requires much support and students are often not able to communicate their reasoning well without scaffolding. One of the most effective ways is to show examples of weak versus strong reasoning (McNeill & Martin, 2011).

Middle school students make claims regularly and often will present evidence to back up their claim, even if it's erroneous. Reasoning provides a feedback loop to themselves to investigate the why behind the claim and ensure that their evidence is solid. Offering a rebuttal phase for students is not often included in most CER activities but can provide feedback for weak versus strong reasoning, errors in claims, misleading or misunderstood evidence, and practice for peers to identify how to argue positively in pursuit of the best explanation behind phenomena (McNeill & Martin, 2011).

Argument Driven Inquiry

Argument driven inquiry (ADI) broadens the CER model to take on more challenging concepts. Students work in small cooperative groups to make a claim in answer to a complex yet specific guiding question. An example of a question might be, "Why do earthquakes usually accompany a volcanic eruption while volcanic eruptions do not occur with every earthquake?" Students are given written text resources, data tables or maps, active scientific websites that monitor the phenomenon in question, models, links to animation and other materials that refer to the guiding question. Working together with instructor approval, students create a plan for the investigation using the materials. Teams to then come up with a claim, an analysis of the data and an explanation of the evidence they find to back up their claim. Their analysis goes through peer review wherein peers critique their work both empirically for quality and sense as well as theoretically for its soundness and usefulness as an explanation. Once arguments are set, students write a lab report style paper to answer the guiding question which also undergoes peer review. Finally, students take their peer-reviewed paper and refine their arguments as needed (Sampson et al., 2009).

The research team behind ADI ran a controlled action research project with a class of 19 high school chemistry students. To determine the efficacy of ADI on student written and oral scientific argumentation, students participated in 15 ADI activities over an 18-week period. The goal was to bolster written and oral scientific argumentation in a student population. Students remained on the same team for the entire time and went through the process weekly. Pre- and post-intervention activities measured students' argumentation skills against norms of accuracy, collaboration, active listening and response and rebuttal. In the beginning, single students dominated the discussion, whether they were correct or not, and the other students tended to confirm what the leader was saying. The claim and evidence given to explain phenomena were inadequate or incorrect and all other aspects relied on informal argumentation rather than scientific theories or laws. Post-treatment, there was a significant change in presenting a correct, more robust claim with corresponding hard evidence. However, students still relied on very general, informal argumentation rather than citing specific laws. Additionally, all students engaged in near-equal amounts of discourse with true argumentation statements such as challenging claims, investigating claims together, responding to fallacies with redirection to correct ideas, and proposing alternate explanations for phenomenon that might warrant investigation to rule out other ideas. Like other instructional techniques, this research reveals that for most students the development of argumentation skills following a scientific model looks to be a progression that takes place over time and improves with practice (Sampson et al., 2010).

Reasoning about Data in Middle School Science

Niels Bohr (1934), the Nobel prize winning physicist and educator, stated "the task of science is to both to extend the range of our experiences and reduce it to order" (p. 1). Early work in inquiry as an instructional practice gave rise to cooperative learning practices such as CER, POGIL, and ADI as forms of inquiry style learning based on student experiences. The National Research Council confirmed these findings in 1995 that the central strategy for teaching science should be inquiry into authentic questions that arise from student experiences (National Research Council, 1995).

Researchers Vellom and Anderson (1999) explored this strategy in a study that created "a rich problem space" for a classroom of urban sixth graders with no prior experience using inquiry in the classroom (p. 184). Researchers created this space by providing open inquiry investigation into liquid density through experience. Students were allowed to create their own investigation to determine what combinations of different colored liquids were possible according to their respective densities. The goals of the study were fourfold: reach consensus on what was not possible in a task, developing persuasive arguments that relied on relevant data while ignoring the irrelevant, getting students to appreciate and understand that science knowledge is a rigorous communal effort that differs from opinion, and finally to establish classroom norms for inquiry conducted as a community (Vellom & Anderson, 1999).

Reaching consensus on a claim involves peer review and rebuttal. It requires students to have solid evidence, to make rigorous claims, and to be able to argue persuasively based on evidence. Therefore, the first two goals were deeply intertwined with one another. The third goal dealt with the nature of science and the fourth with identifying science and engineering processes as well as the nature of science (Vellom & Anderson, 1999).

Another goal of the study was to identify what types of arguments students would make and the veracity of the arguments. Researchers also wanted to observe what students gained as they navigated the science content as well as the practices and processes used. Finally, social interactions within groups were closely observed as students were required to collaborate without assigned roles or guides (Vellom & Anderson, 1999).

Outcomes were varied. Students were able to eventually reach consensus and distinguish between possible and not possible variations on the task. Students also reported a higher level of confidence after having figured it out for themselves. Concerning the second goal, students were successful in presenting reasoning based on personal experience and employed solid rhetoric to argue for their own findings but did not apply larger patterns to the classroom as a whole. The third goal results were much more troubling. Students were able to understand that science is collaborative and require claims to be supported by durable and replicable evidence, but their findings from their task were not seen as evidence of a larger law or theory that applied in other situations they encountered. Finally, students were not able to create a culture of social norms on their own in which all voices were respected and heard. Coercion was more the order than consensus, which required more teacher intervention than was desired and likely led to a deleterious effect for some student voices. In retrospect, taking time before the task to create and explore norms as a class would probably have given low performing students more opportunity. As an early research study into the dynamics of inquiry and student experiences, this was not necessarily a setback without value. The inability for students to create a respectful yet argumentation-friendly environment on their own led to awareness of need to help students, whether through scaffolding in cooperative learning roles or in taking time to cultivate norms through whole group consensus prior to tasks (Vellom & Anderson, 1999).

A Learning Progression for Scientific Argumentation and Reasoning

While one of the central goals for science education is to develop argumentation skills, students clearly need help and scaffolding to develop those skills. They benefit from modeling and clear examples of good reasoning in both content and structure. Teachers need to understand where students are in terms of their progress and how to provide support via the classroom environment and instructional context in which the learning takes place (Berland & McNeill, 2010).

Berland and McNeill (2010) define learning progressions as "pathways to support student learning" or "increasing levels of complexity in disciplinary knowledge and practices" (p. 767). Berland and McNeill identify a learning progression that begins at a basic level and moves with support to enabling students to develop complex argumentation skills. In their research they first identified that students often needed to see examples of good reasoning contrasted with poor argumentation in order to improve their own work. In their comprehensive survey of elementary, middle school, and high school students, three aspects of support were identified as benefitting students of all ages: content, developmentally appropriate data sets, and learning scaffolds (Berland & McNeill, 2010).

Instructional content is the core of most science classrooms and sets expectations for what concepts students should be learning and retaining. It is also the most significant in this study in terms of helping students develop complex argumentation skills. The most effective support in instructional content comes from providing an abundance of resources and experiences from students to draw on rather than a single or few interactions with an idea. When students experience a phenomenon or idea in multiple ways, they can grasp its key concepts better and home in on evidence to respond to questions with legitimate claims. Instructional content needs to provide questions of increasing complexity where the possible claims may vary widely. Interestingly, questions that are simpler often produce more complex and varied argumentation (Berland & McNeill, 2010). Additionally, data sets provided to students should be appropriate to their skill and increase in complexity at pace with proficiency in argumentation. Finally, scaffolds will help students define their argument better. Scaffolding can look like many tools: visual organizers, flowcharts, organizers with questions about how students choose and analyze data, peer critique and many more that would be specific to instructional contexts. Students simply need a framework to build on moving from more to less in terms of support as proficiency increases (Berland & McNeill, 2010.)

The second part of a learning progression which is measured is the argumentation product. An argumentation product can be written or spoken. A good argument contains reasonable support for a claim and has a counterclaim as rebuttal that has been found lacking in support. The third part of a learning progression is the argumentative process. It makes sense that this is closely interwound with the product previously mentioned as it is the method used to gain the claim. In the argumentation process, students should state, defend, evaluate and revise their own claims while questioning and evaluating the claims of others in a back-and-forth manner. Students naturally do not engage in challenging the ideas of others due to social tension, so providing scaffolds to allow argumentation as an expectation is important (Berland & McNeill, 2010).

CHAPTER THREE

METHODOLOGY

POGIL is designed to introduce or extend concepts while ADI is typically used as part of a larger unit to give students an opportunity to apply knowledge in argumentation. Therefore, I wanted to examine the effects of using a POGIL activity designed specifically to investigate the components of claim, evidence and reasoning and use various instruments to assess students progress through treatment. Students would undergo two rounds of ADI before and after intervention and their argumentation products would be evaluated.

Treatment

A background knowledge probe, pre- and post-tests, surveys and open-response questions were used to collect data both before and after treatment during this action research project.

The intervention strategy was applied to two groups of sixth and seventh graders using a POGIL that explored the argumentation components of the Claim, Evidence and Reasoning (CER) framework. The research methodology for this project received an exemption by Montana State University's Institutional Review Board and compliance for work with human subjects was maintained (Appendix A). Ten of my eighth-grade students piloted the first iterations of the instruments and treatment along with the assessment. This study was designed to investigate if addition of a POGIL activity specifically designed to help students investigate the components of CER would make a difference in being able to identify valid claims, evidence, and reasoning as well as improve proficiency in student abilities in writing their own claim, evidence and reasoning statements in response to a specific query such as an argument-driven investigation (Appendix B).

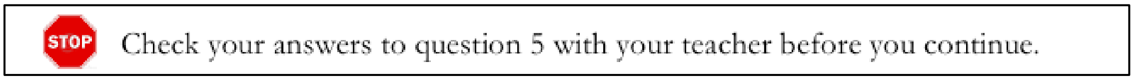
The treatment group consisted of 18 sixth grade integrated science students and 18 seventh grade integrated science students. The seventh-grade class had some exposure to CER instruction and practice the previous year, but our state's pandemic response had shut down our school for in-person instruction and ended our school year with less-than-ideal instruction opportunities during the last quarter. Additionally, there were seven new students in seventh grade who did not attend our school last year and who did not indicate a high familiarity with the nature of science argumentation. Sixth grade had very little science instruction previous, mainly by me in half-hour enrichment classes that met once weekly for part of a semester. Thus, both groups could stand to benefit from targeting scientific argumentation skills.

Since this research involved science process skills that reach across content, content throughout the pre-treatment, treatment and post-treatment phase included a life science unit and an Earth science unit. The ADI life science investigation took place after a unit on human body systems and incorporated the NGSS disciplinary core idea relating to structure and function of cells in an organism (MS LSA.1). More specifically, students would contrast animal cells with plant cells and seek to identify photographs of unknown cells that had large and obvious features which animal cells do not generally have, such as large vacuoles, cell walls and chloroplasts. The ADI earth science activity took place after a mineralogy unit; students used a set of known minerals which they tested to identify a set of unknown minerals with those same qualities. Looking at patterns in mineral structures through physical properties builds background knowledge about how Earth's materials cycle (MS ESS 2).

Prior to completing ADI activities, students were given a background knowledge probe and a pre-test where they had to identify statements as claim, evidence or reasoning. They also took a survey about their perceived proficiency in argumentation. This phase took place during the week of December 7-11, 2020 and included the first attempt at ADI. These three instruments were typical for tasks I would use in class.

Before any exposure to CER inquiry or new instruction, both groups attempted an argument driven inquiry activity that required them to present a supported claim identifying plant cells and animal cells from unidentified slide photos. They created group posters for their claims. Each group received feedback and critique on their claims, evidence and reasoning in a poster session from other students before turning in a personal version of a claim with evidence to support it and reasoning to explain. This allowed for individuals to disagree or improve on their group's work. My fellow science teacher Mare Sullivan then scored the individual responses using a CER scoring rubric.

Next, I began the treatment phase. The treatment protocol consisted of a POGIL activity designed specifically to examine the components of claim, evidence, and reasoning. Specifically, the activity contrasted the differences between each component and examined how those components related to each other in scientific argumentation. Like many POGILs, it begins by presenting a realistic scenario. This particular scenario includes arguing for and against a claim regarding a problem and uses tables with models, data sets and data collection exercises to establish evidence. Students worked with evidence to evaluate claim validity between three choices, then collected further evidence and explored reasoning that could tie their evidence to one of those three possible claims. Group members used defined roles of reader, manager, and quality controller over two to three days using this format. The final piece of the guided inquiry activity is the creation of a claim, evidence and reasoning section that also includes rebuttal. During the treatment activity, I roamed between groups, checking in at predetermined points marked by a symbol (a stop sign) to indicate when students should stop or check in before continuing. This allowed me to give instant feedback and re-direct any misconceptions as well as address any comprehension gaps or other questions that came up. Students also reported out to a class idea chart on certain questions and were allowed to send envoys to double check their ideas against other groups (Figure 1).



Check your answers to question 5 with your teacher before you continue.

Figure 1. POGIL stop sign.

Finally, students engaged in a second round of argumentation using a second ADI lab activity. This time the guiding question concerned identifying unknown minerals by comparing properties of known minerals to the mystery set. They gathered evidence, made a claim, and presented reasoning to connect their claim to evidence. Posters, critiques and feedback were given once more by peers, then each student was responsible to take their group's claim, evidence and reasoning and put it in their own google slide that followed a set format (figure 6). If they disagreed or wanted to add more information in any category, they could change it for the better, but they had to minimally include their group's work.

Instrument Description

To begin the action research process, I first administered the Background Knowledge Probe. This is a form of classroom assessment technique that allows an instructor to gauge the level of familiarity with a topic before study. While background knowledge probes can be more complex if desired, I simply wanted to know if my students remembered or had experience with the claim, evidence and reasoning framework and so only that question was asked. Students were given 5 minutes to respond. All students were present to respond on a paper copy, but I made a copy available for distance learning as our school had its first COVID cases the first week in December and some classes were sent home. My original plan for the probe response was to rate each student's perceived proficiency along a scale of 1-3. A student showing little to no familiarity with claim, evidence and reasoning would have been given a score of 1, a student who remembered using it or had experience with it in a previous school had a score of 2, and a student who was able to explain most or all of it would have received a score of 3. I originally planned to use those scores to create small POGIL and ADI groups with a mixture of self-identified high, medium and low familiarity students (Appendix C).

What I didn't anticipate was the diversity in answers I would find in the probe. The answers were so widely discrepant that my plan was scrapped in favor of putting one student in each group who was comfortable with CER and using classroom observations about who worked well together as a team. This ended up working well and allowed me to change the groups at a future date in case of issues in personality differences or work style.

To collect information about student attitudes regarding proficiency, I created a Student Attitudes Regarding Claim, Evidence and Reasoning Survey to determine how students felt in using different aspects of claim, evidence and reasoning. I translated this survey into a google form as well in anticipation that we would have some students doing distance learning due to COVID exposure; two students did have to use that method to respond. Students were given statements about identifying, creating and connecting CER components. They were then asked to choose a response of strongly agree, agree, disagree or strongly disagree for each of the six statements. A seventh item at the end of the survey invited students to share whatever else they felt I needed to know. The survey was given pre- and post-treatment and analyzed for change (Appendix D).

One of my guiding questions prompting this action research project was whether or not students could identify claim, evidence and reasoning statements. I designed the Identifying Claim, Evidence and Reasoning Statements pre-test to evaluate that proficiency. The twelve items on the test ranged in subject matter from sports, plant versus animal taxonomy, and extended comically to the possibility that a student could be a zombie. Several ideas were connectable to one another in the hope that students would be able to use patterns and reasoning to sort these ideas into supportable claims, observable evidence and connected reasoning. The ideas were randomized and placed into a grid. Students could indicate one of three choices by writing C, E, or R beside each statement to correspond to claim, evidence and reasoning. This was also available as a google form for students who were under COVID quarantine, but all who took it were able to use the paper forms (Appendix E).

Their answers were evaluated for accuracy and data was collected. After the treatment period, students were given the same test again and the pre/post-test scores were compared were using normalized gains. The normalized gain was calculated to determine differences in student knowledge. Normalized gains of less than 0.3 percent were considered low, gains 0.3 to 0.7 were considered a medium gain and normalized gains greater than 0.7 were considered high (Hake, 1998.)

Another goal for treatment was the desire to improve CER statements made by students on a continuum of accuracy, completeness and logic. This meant a rubric of some sort would need to be present for students to have a tangible goal. A rubric was used to evaluate claim, evidence and reasoning statements made by students both before and after treatment. While there are many rubrics available for CER writing, I chose one that seemed straightforward to students and simple to score but gave more weight to the reasoning portion (Meacham, 2017). The rubric gave clear explanations of each type of statement and expectations for scores. A zero rating reflected an inaccurate or missing claim, inappropriate or missing evidence, and non-supportive or missing evidence. Statements receiving a 1 rating on claims were accurate but vague or incomplete, while evidence in this rating group had to be mostly appropriate but might include some errors. Reasoning in the mid-range repeated the evidence with some connection to scientific principles and received a weight of 2 points. The highest rating for claim required an accurate and complete claim to receive a weight of 2 points, while the top evidence rating had an identical score and necessitated evidence presentation that was both appropriate and sufficient. The highest number of points a student could receive for reasoning was 4. Reasoning at that level required accurate and complete logic that connected the claim to the evidence using appropriate and sufficient scientific principles (Figure 5). Since reasoning demands more critical thinking than making a claim or gathering data for evidence, it was hoped that a heftier point value would help students remember to give more effort to that piece of the framework. Changes in argument proficiency were analyzed using normalized gains as well.

Due to Covid restrictions, face-to-face interviews were difficult to conduct so I asked the following question on the student attitude survey: "What else would you like for me to know?" While many students chose to use that space to give me feedback on how their ADI groups had worked or not worked well, several students responded specifically to learning how to approach CER. Additionally, the ADI lab had checkout questions that I used to evaluate students individually. Accurate answers reflected appropriate reasoning.

In working with writing CER, one aim of the study was to evaluate if students would demonstrate an increase in their ability to differentiate between information that could be considered evidence and information that could not support their claim.
I used student written CER responses both before and after treatment to evaluate this question. Only responses with evidence that did not support their claim were counted.

Additionally, the student attitude survey and open-ended questions gave opportunity for students to reflect on their perceived ability to identify supporting evidence for claims.

Instrument List and Data Triangulation Matrix

Focus question and sub-questions	**Instrument and Methodology**	**Instrument and Methodology**	**Instrument and Methodology**	**Instrument and Methodology**
How will using POGIL affect student proficiency in making their own claim, evidence and reasoning statements?	Likert Survey questions regarding Student Attitudes Pre-and Post - Treatment	Background knowledge probe	ADI CER evaluated through common rubric	Student open response question
Sub-question: Will student abilities to discriminate between claim, evidence and reasoning statements improve as a result of POGIL?	Pre-test and post-test of CER statements	ADI CER evaluated through common rubric	Likert Survey questions regarding Student Attitudes Pre- and Post - Treatment	Student open response question
Sub-question: Will student ability to identify evidence supporting a claim versus evidence that refutes or has no correlation with a claim improve as a result of POGIL?	Pre-test and post-test of CER statements	ADI CER evaluated through common rubric	Likert Survey questions regarding Student Attitudes Pre- and Post - Treatment	Student open response question

CHAPTER FOUR

DATA ANALYSIS

Results

The results of the Student Attitudes Regarding CER survey indicated that 82% of sixth grade and 76% of seventh grade students who chose to respond agreed with the first question (N=33). Both classes indicated a similarly high confidence in being able to identify a claim statement for the second question. 88% of sixth grade and 95% of seventh grade students indicated agreement to strong agreement in being able to identify evidence to support a claim per the third item on the survey. However, when asked about differentiating between evidence and reasoning in the fourth question, 25% of the combined classes disagreed or strongly disagreed. A similar drop in confidence showed up in response to question five. Agreement was high concerning the statement about student ability to use reason to explain why evidence supported a claim on item six. However, some students in both grades strongly disagreed that they were able to achieve this goal.

Post-treatment, students were given a chance to respond to the same survey regarding their personal growth and perceived aptitude towards CER argumentation. Overall students felt more confident across most areas except for being able to assess whether evidence supports a claim (Figures 2 & 3). That measure decreased overall, but none of the students disagreed strongly after treatment.

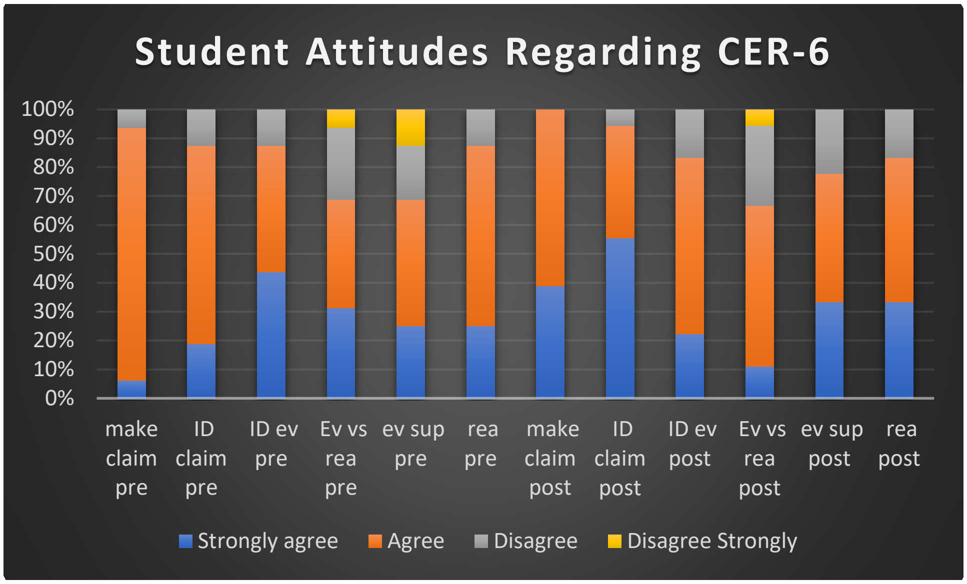


Figure 2. Student attitudes pre- and post-treatment regarding CER argumentation ability, grade 6, (*N*=16).

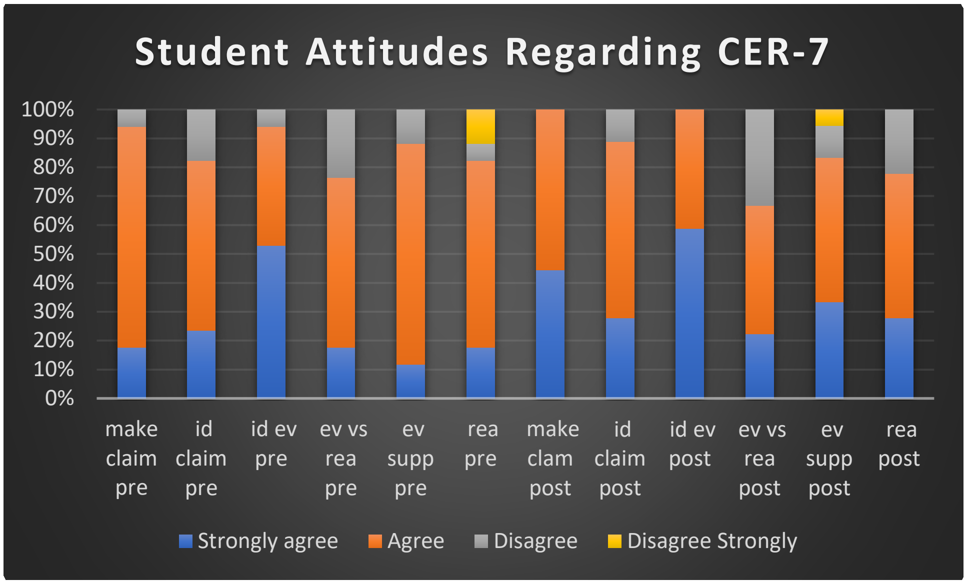


Figure 3. Student attitudes pre- and post-treatment regarding CER argumentation ability, grade 7, (*N*=17).

Results of the pretest for Identifying Claim, Evidence and Reasoning Statements were 69% for seventh and 66% for sixth grade (*N*=36). None of the seventh-grade class scored null on the pretest; one of the sixth-grade students could not identify any statement correctly. Posttest results were 77% and 75% respectively. Pre- and post-test scores over the ability to identify claim, evidence and reasoning statements were compared using normalized gains. Seventh showed a normalized gain of .28. Sixth grade showed a normalized gain of .27 (Figure 4) both of which are low gains according to Hake (1998).

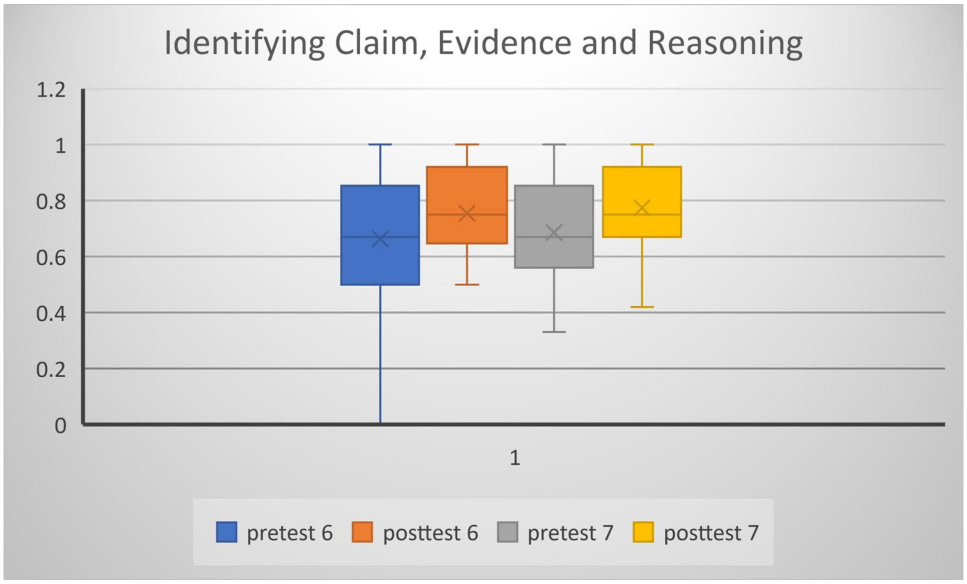


Figure 4. Gains from identifying claim, evidence and reasoning pre-and posttest for grades six and seven (*n*=18, 6th grade; *n*=18, 7th grade).

Students engaged in inquiry in response to a guiding question and submitted the results of their investigation by making a claim, presenting their evidence and attempting to demonstrate reasoning to connect their evidence to their claim. This was done pre- and post-treatment (*N*=35). Student work was then scored using the CER Scoring Rubric

(Appendix F). When work was evaluated using normalized gains, seventh and sixth showed low gains in presenting argumentation (Figures 6 and 7).

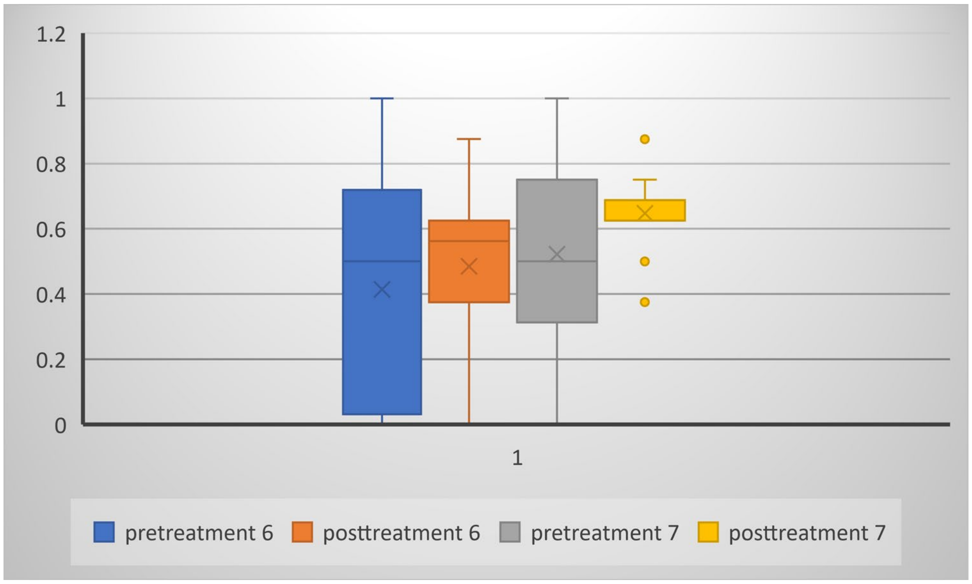


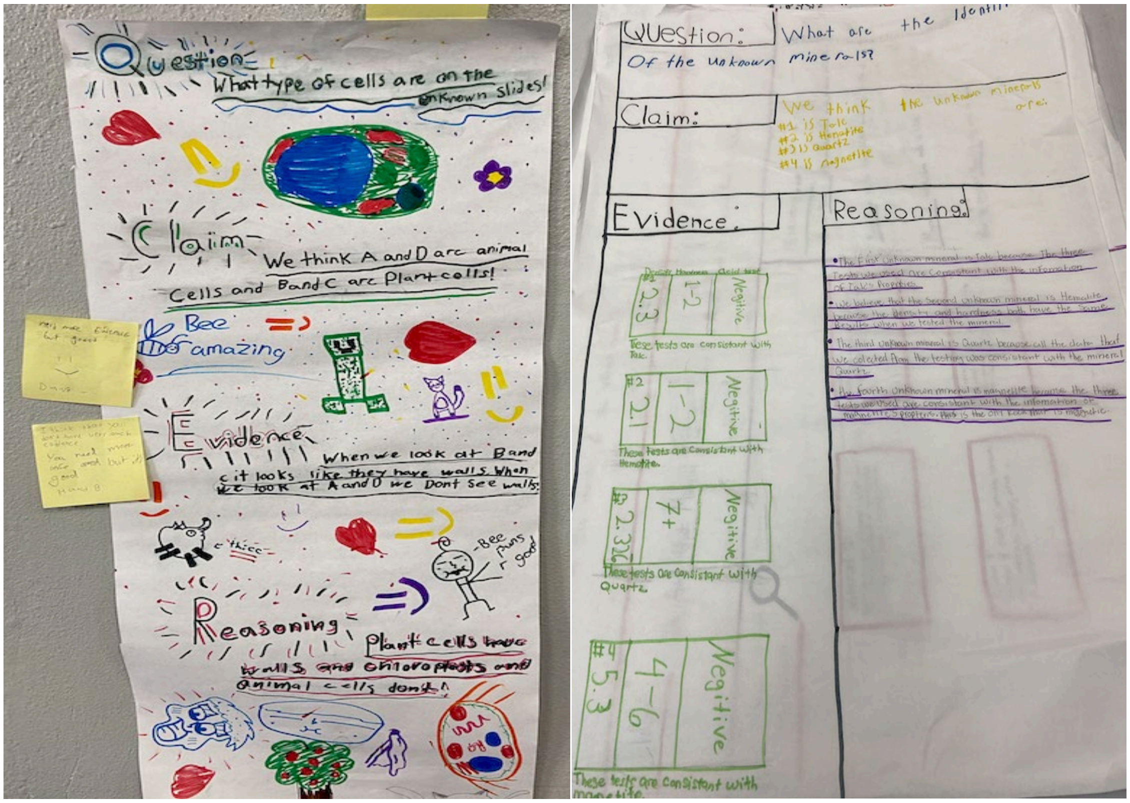
Figure 5. Gains in writing CER pre- and post-treatment for grades six and seven (*n*=18, 6th grade; *n*=17, 7th grade).

In working with writing CER, one aim of the study was to evaluate if students would demonstrate an increase in their ability to differentiate between information that could be considered evidence and information that could not support their claim. Written CER arguments were evaluated for claim appropriateness. Pretreatment, 43% of sixth graders and 41% of seventh graders did not use appropriate evidence to support their claim. Posttreatment, only 13% of sixth graders continued to include inappropriate evidence and none of the seventh graders included evidence that did not correlate to the claim they made. Convention in language and data organization had improved as well with more complex data and tables as well as attempts at using tests and techniques that would give them quantitative data. Students also reported they felt more confident in being able to identify evidence supporting their claim; 77% of sixth graders and 82% of seventh graders agreed or strongly agreed that they could identify evidence that supporting their claim.

Covid restrictions meant that we needed to take precautions to protect both students and teachers from possible infection. Our area saw an increase in cases just as we finished our initial argumentation, so I chose to use the final question on the student survey in lieu of student interviews as well as a question on the post-treatment ADI lab checkout. One sixth grader wanted me to know that he felt like he could do CER now, and his posttest score was improved to corroborate that attitude. Meanwhile, a student who typically carries an A average and is a diligent student wrote, "I think I need to work more on this topic (CER)". She specifically stated that she often flips reasoning and evidence and has trouble telling where one stops and the other starts. One of my students who often receives one-to-one academic support stated that in order to do CER you first need to have a guiding question and remember it during your investigation. Finally, one sixth grade student shared that she still felt like she had a hard time differentiating between evidence and reasoning, but her post-test scores were improved while one of her more self-described confident classmates did not see significant increase.

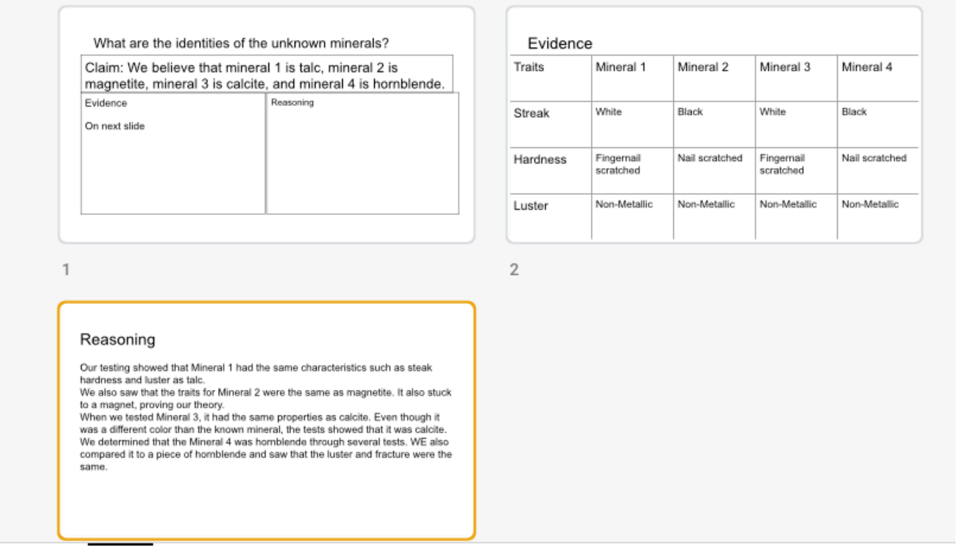
ADI labs typically include post-activity questions which require students to use reasoning and evidence to explain their choices. The checkout question from the initial lab had two final questions relating to the importance of identifying patterns in science along with how the structure of an organisms affects its function. Both questions asked students to relate their investigations to these concepts (Enderle et al, 2015). Students did not address either question appropriately and either left it blank or explained what patterns they found. There was a significant change on the second checkout question set post-treatment. When asked why it was possible to identify a mineral based on physical properties, 22% of seventh grade students were able to state that this is a possible task because minerals have unique properties that can be identified using physical tests and 33% of seventh grade students were able to state that clearly.

Student-created CER slides and posters changed from the beginning to end. The pre-treatment posters were decorative with minimal text, while the post-treatment set were heavy with organized data in the evidence section (Figure 6). Initially, many students restated their evidence or claim in lieu of reasoning. Reasoning on post-treatment posters and slides had more text, although the majority still kept including evidence in their reasoning. On their post treatment personal responses, several students were not content to use the space provided and created a second slide to create space for evidence tables and reasoning (Figure 7).



.

Figure 6. Examples of initial (left) and final (right) argumentation posters.



What are the identities of the unknown minerals?

Claim: We believe that mineral 1 is talc, mineral 2 is magnetite, mineral 3 is calcite, and mineral 4 is hornblende.

Evidence	Reasoning
On next slide	

1

Evidence

Traits	Mineral 1	Mineral 2	Mineral 3	Mineral 4
Streak	White	Black	White	Black
Hardness	Fingernail scratched	Nail scratched	Fingernail scratched	Nail scratched
Luster	Non-Metallic	Non-Metallic	Non-Metallic	Non-Metallic

2

Reasoning

Our testing showed that Mineral 1 had the same characteristics such as steak hardness and luster as talc.
We also saw that the traits for Mineral 2 were the same as magnetite. It also stuck to a magnet, proving our theory.
When we tested Mineral 3, it had the same properties as calcite. Even though it was a different color than the known mineral, the tests showed that it was calcite.
We determined that the Mineral 4 was hornblende through several tests. WE also compared it to a piece of hornblende and saw that the luster and fracture were the same.

Figure 7. Example of final student argumentation.

CHAPTER FIVE

CLAIM, EVIDENCE AND REASONING

Claims From the Study

When I began designing the study in anticipation of conducting action research, I knew I wanted to work with targeting the process skill of argumentation. My middle school students made acceptable claims, showed widely varied ability across grades in gathering evidence, and demonstrated very little proficiency in reasoning skills. It seemed like instruction and inquiry in using argumentation only slightly improved their ability to connect claims to evidence but had little to no impact on reasoning. Using a specifically designed POGIL to address this gap brought specific, clear instruction to this group and allowed them to investigate the differences between the three main aspects of argumentation. The whole process of action research brought a spotlight to skills that needed work and raised student awareness that it was an important process skill for science and many other applications as well. Responding to a background probe, taking pre and post-tests to evaluate their ability to identify CER statements, and working through a POGIL as a group that allowed them to break down the elements of CER all contributed to focusing on this particular process skill.

While neither ADI nor POGIL are defined as learning progressions, both provided guided argumentation experiences from different approaches. One can see POGIL as reducing it to order and ADI as extending its reach. By using POGIL as a scaffold, students had opportunity to extend proficiency in argumentation. By engaging in ADI through using the skills gained through POGIL scaffolding, students exhibited progression by using guiding questions, planning investigations with teacher feedback, and presenting arguments for critique.

After all the data was collected and analyzed, students overall showed gains in their ability to write CER, their ability to discriminate between statements as claim, evidence or reasoning statements, and their ability to select appropriate evidence to support a claim. From a qualitative perspective, their argumentation became much richer and more organized. It is my claim that adding a guided inquiry activity to our learning environment brought improvement by giving them a better scaffold.

As an instructor conducting action research for the first time, I found the process rich and informative as I examined student work for patterns and documented changes to student understanding. It was enlightening to look at background knowledge probes which asked students to explain their ideas about claim, evidence and reasoning and see the diversity in understanding. One student thought he'd be great at CER because he "watched a lot of cop shows." Several students described claim as indicating possession or discovery; others were able to specify that claim means to state that something is true and prove it. Students were more on target with evidence in terms of definition. The word proof was used by most students in relation to evidence. However, by the end of the study, most students used the word claim appropriately as evidenced on their work through investigation and were able to exclude data that did not support their claim.

Reasoning was the most difficult thing for students to describe, with evidence being restated alongside variations of the words "logic" and "explanation" by a few students. This affirmed what I saw in their work after all treatment was concluded; in spite of good examples and clear instruction and experience, meaningful reasoning remained elusive for some students.

This raised more questions for me. Is it developmentally appropriate to approach reasoning with younger adolescents? Gains in these areas give me hope that it isn't unobtainable, but proficiency might be related to exposure and experience. Refutation or rebuttal kept coming up in conversations with students in concept if not in name. An open question remains for me as a teacher, then, about using refutation as a tool to improve reasoning.

Action research is by nature never complete, and I believe as a beginning practitioner there were many directions I could have taken to examine what was happening when I targeted argumentation through the use of POGIL. Like any new endeavor, many of the lessons are found when problems are encountered, and my action research project was no exception.

It was an unusual year due to COVID protocols and that restricted many design choices. I had a shorter instructional week and students were limited in group work because of social distancing. Additionally, we had many students out for long absences due to quarantine or illness. I would have preferred to compare a treatment group with a non-treatment group. I was concerned that it wouldn't be an accurate way to measure progress due to small class sizes, but mostly I wanted to make sure that each student had access to what I hope would be better learning opportunities through the inclusion of POGIL.

My eighth-grade class beta-tested every instrument, the treatment POGIL and the argumentation investigations. They alerted me errors or unclear word choices. While that class growth in argumentation seemed significant, their class population nearly doubled throughout the year which changed the dynamics considerably and made comparison difficult to navigate. In a typical year, I would have liked to have taken their measure as well and included them in the study. I wondered if their growth as older students would have been more significant since most of them had more opportunities to engage argumentation activities. These are just a few things I would do differently.

Part of the research that surprised me was how deeply I enjoyed creating databases and working with the data to uncover bigger patterns. It was enlightening to give shape quantitatively to things I had informally observed in students in class; sometimes there were new insights into thinking patterns or misconceptions while other observations were confirmed empirically. It was satisfying to look for the story in the data and read the ending in the gains, giving me insight in what to work on next year. Within the data itself, there were subsets of information such as individual gains or gains as a group on specific items which each tell another piece of the story of what my students gained this year in argumentation skills.

Value of the Study and Consideration for Future Research

The value of my study lies in uncovering misconceptions students have about argumentation, observing what changes are possible when the process skill of argumentation is targeted for treatment and practice, and learning what still needs improvement in this area for middle school students

Additionally, my small study is part of a larger, growing body of research that looks at how students grow in reasoning ability and what types of pedagogical choices can be made in the classroom to help them in this endeavor. I found value in close examination of the effects of guided inquiry towards a specific process skill and I would hope others could benefit from this study as well.

Future research could entail continuing to teach process skills within content area coordinating with refutation and refinement of reasoning. There were many more questions I had about student learning processes and research into any of them will yield information to help me improve as a teacher. General background vocabulary in science is not typically high in middle school and students often had a difficult time finding words to express their ideas. I wonder what impact deliberately building meaningful science literacy would have on argumentation. Finally, after watching my students learn the value of creating tables and charts on their second ADI investigation,
I would like to explore more about building meaningful, student-generated data sets that students can explore as evidence and look for reasoning within those patterns.

Impact of Action Research on the Author

At the beginning of my time in the MSSE program, I was concerned about conducting an action research project successfully. After going through the process, it has been one of the most informative years of my teaching career. I've become much more deliberate about what we do and how I teach concepts rather than merely using the same approach I used previously. Every group of students is unique and requires careful assessment of their needs with a response that is appropriate. I would hope to continue using many types of the instruments I used this year to uncover areas where students are needing scaffolding, where learning progressions might be implemented, and how I can help students learn to communicate through written argumentation at a more coherent and cohesive level. Through using action research processes, I can see the impact of instruction on a much more profound level rather than just looking at success or failure through grades and mastery of concept on tests.

This means that I will have to work on communicating new ways of assessment and evaluation to students and parents. Our school has typically not used standards-based grading or problem- based learning approaches but instead used a more traditional grading system, but I can work within that system to reflect growth and change. Gradual instructional changes over the past nine years have resulted in a science teaching program that looks very different than it once did, so more changes won't be too shocking but will require adjustment.

CHAPTER SIX

THE JOURNAL ARTICLE AS SUBMITTED TO SCIENCE SCOPE

Title: Learning to Argue

Subtitle: Using Process Oriented Guided Inquiry Learning to Increase Student Proficiency in Claims, Evidence and Reasoning

Abstract: The author's classroom action research project's purpose was to examine the effects of an educational approach called Process-Oriented Guided Inquiry Learning intervention on student abilities to create, identify, and understand different aspects of scientific argumentation. The project follows two middle school classes through the process of pre-and post-assessment, treatment, and inquiry labs where they engaged in argumentation in response to a guiding question. Students in grades six and seven investigated the differences among claim, evidence, and reasoning in scientific argumentation. The research showed gains in each group as measured through identification of CER statements and writing their own CER argumentation.

Inquiry is a strong tool; it has been the primary vehicle recommended for science instruction in the United States for over 25 years (National Research Council, 1995). The NRC further clarified that the current three-dimensional model of science and engineering practices, cross-cutting concepts, and disciplinary core ideas is not meant to replace inquiry but communicate a deeper understanding of how to teach it (National Research Council, 2012). When students engage in inquiry, their findings are communicated through argumentation, often in the form of claim, evidence and reasoning

(CER). This approach models how many scientists conduct research and provides an opportunity for students to think like scientists and evaluate others' claims.

In middle school science classrooms, the NGSS disciplinary practices of planning and carrying out investigations, analyzing and interpreting data, and engaging in argument from evidence are critical steps towards conducting effective inquiry. As a science teacher for grades 6-8, I consistently observed that my students had little to no background in these practices when they arrived at my science class in grade six.

Investigation opportunities provided good initial experiences. Argument Driven Inquiry activities were implemented a few times per year alongside an inquiry-style science fair. Students were soon able to identify accurate claims but found analyzing and interpreting evidence in the form of data to be challenging. They struggled to plan or carry out independent investigations even when given a guiding question.

I began to search for ways to potentially foster gains in argumentation skills and wondered if breaking down the component parts of CER be helpful to students and improve their argumentation skills. As a student in the Master of Science in Science Education program at Montana State University, I decided to undertake an action research project to see if a targeted intervention strategy would be effective. Dr. John Graves, my advisor, worked with me through MSU's foundation classes to help me design the project.

I wanted to use a tool that learners already knew how to use and found enjoyable. Process-oriented guided inquiry learning (POGIL)was familiar to students, and it ranked among my students as one of their favorite learning tools on end-of-year surveys.

POGIL is a student-centered approach that uses small teams with assigned roles to explore concepts. Models are used to explore aspects of fundamental ideas, learn terms associated with concepts, and apply concepts to more complex problems. Students are assigned or choose roles that involve participation through reading, timekeeping, documenting, and collaborating with the larger class.

While POGIL focuses on several process skills to develop alongside content in the science classroom, I was most interested in working with my students to enhance written communication, problem solving, critical thinking, and information processing (Moog, 2019). After digging into research on ADI, argumentation, and POGIL, I felt I had a good intervention to try. I've helped developed POGILS for several collections, so I decided to use its structure to work through argumentation components.

With help from fellow POGIL practitioners, an activity that explored claim, evidence and reasoning was born from an investigation that occurred in my classroom some years before. The title of the POGIL was "Is it Safe to Use Your Lunchbox as a Bowl?" and explored the question of food safety and cleanliness in relation to levels of bacteria and dirt in student lunch containers. Students were presented with an inquiry project's data and worked through queries about what they were observing in collected data and experiment results. The final task was to create a claim, evidence, reasoning, and refutation model that supported a claim that argued for or against eating directly from one's lunch box surface.

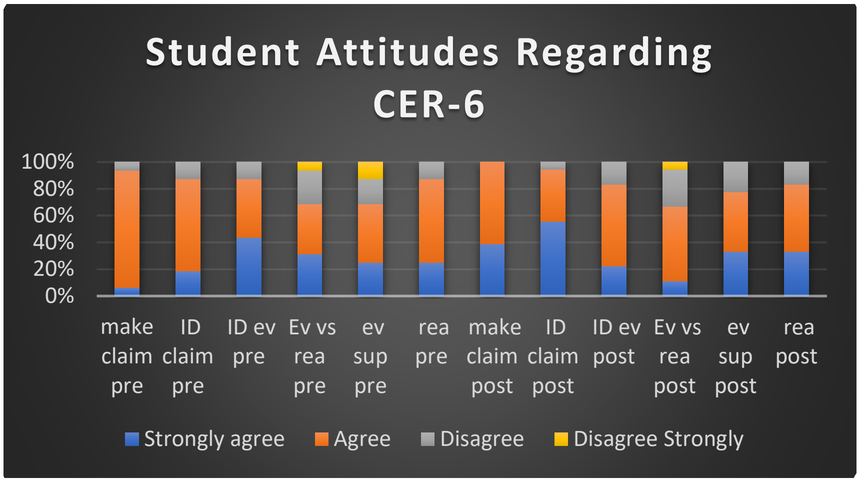
Surveys were given to students to explore attitudes before treatment. A high percentage of students (69% of sixth graders, 77% of seventh graders, N=33) agreed or strongly agreed that they were able to differentiate between evidence for a claim and reasoning that supported a claim. Only 6% of sixth graders strongly disagreed while no seventh graders showed negative attitudes about proficiency.

Next, a series of claim, evidence or reasoning statements were given to students to identify. Sixth grade averaged 65% and seventh averaged 68% on their ability to identify CER statements.

Prior to the treatment POGIL, students engaged in their first ADI unit which generated both group and individual responses; these were scored using a rubric. The initial CER products were diverse; some students unable to create any argument while few approached proficiency. Most students did not meet the standard. The group argumentation posters were scant on evidence, inaccurate on claims, and had little reasoning. Students as individuals likewise struggled to make a valid claim, presented disorganized evidence, and often restated evidence in place of reasoning. The scores averaged 2.94/8 for sixth grade and 3.94/8 for seventh.

After the initial round of inquiry, POGIL groups were organized to provide diversity and support. Students worked through the models and answered questions. I facilitated their discussions and checked for understanding. Students were given claim choices to match their evidence but were not provided reasoning options. Most groups easily discovered an effective claim and were able to provide evidence. Disagreement popped up in groups in writing the reasoning portion.

The POGIL intervention was followed with a second round of ADI, CER identification test, and attitude survey. Students showed gains on both the argumentation rubric and CER post-test. An interesting shift in seventh grade student attitudes towards perceived proficiency communicated needing growth in differentiating between evidence and reasoning.



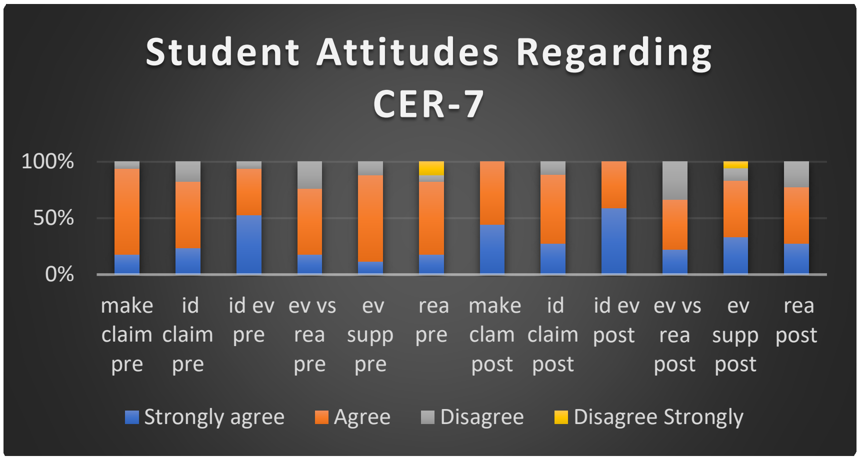
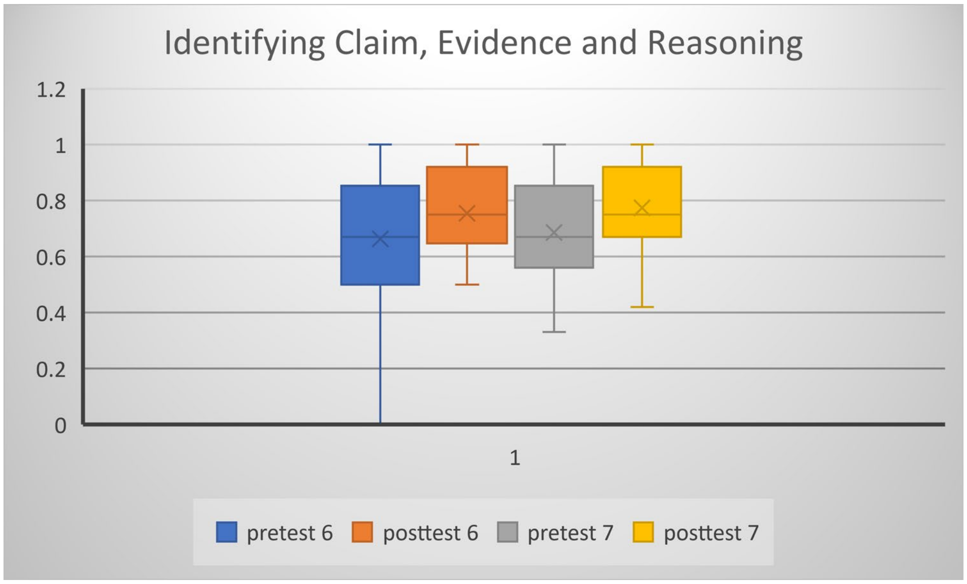


Figure 1. Student attitudes pre- and post-treatment regarding CER argumentation ability, grade 6 (top) and grade 7 (bottom). (*N*=33, *n*=16, *n*=17).



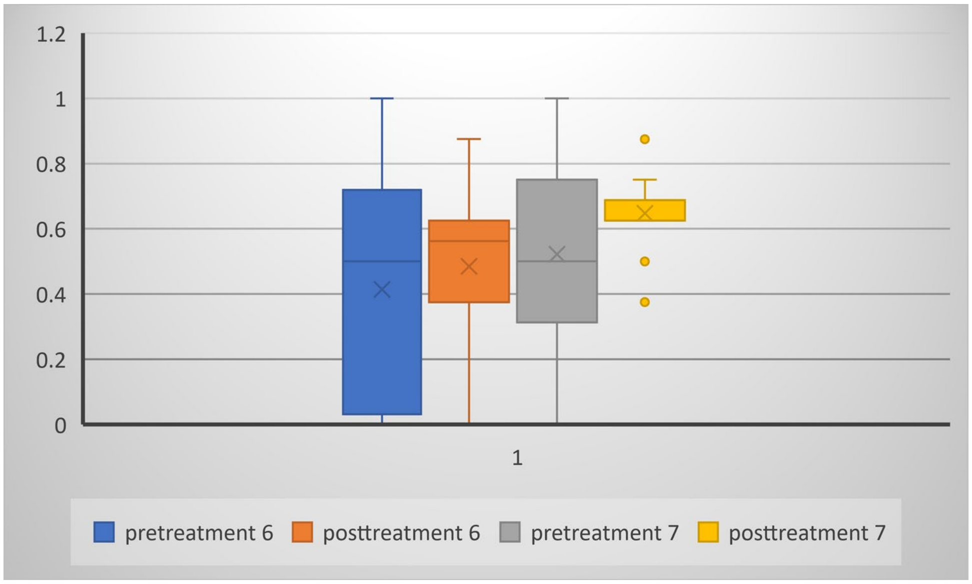


Figure 2. Identifying Claim, Evidence and Reasoning Statements (top) and Writing Claim, Evidence and Reasoning (bottom) (*N*=36; *n*=18, *n*=18).

Qualitatively, the biggest change came in communicating results. Posters went from disorganized with sparse evidence to elaborate and organized data tables filled with measurements in support of accurate claims.

Finally, one measure showed surprising growth. Lab checkout questions accompany ADI investigations which I used to check for understanding. The first question prompted students to provide reasoning but no student from the first round responded accurately. After the treatment, 30% of the students were able to respond with adequate or proficient reasoning statements.

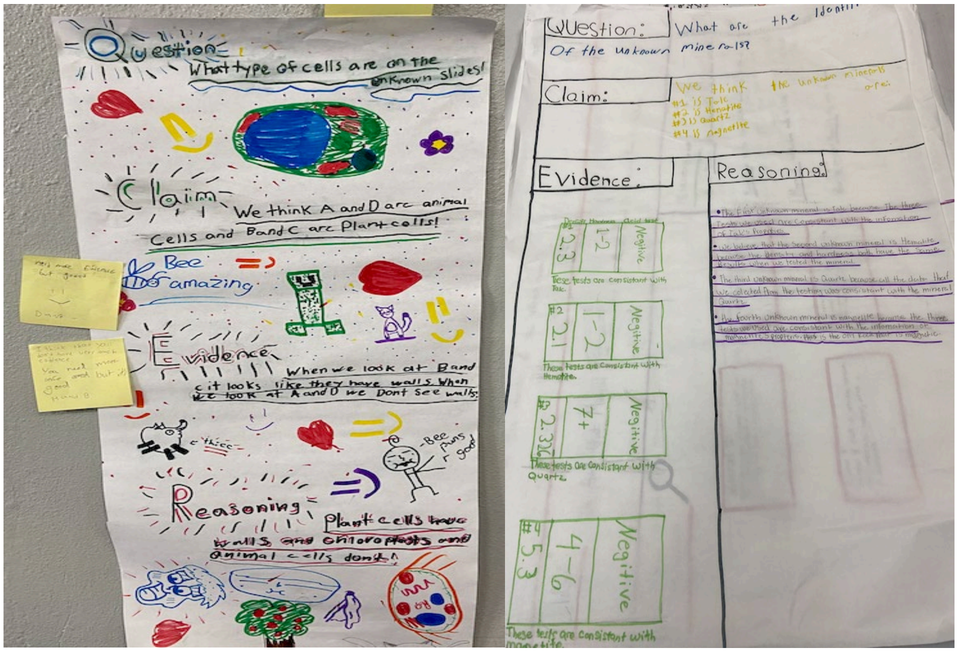


Figure 3: Group CER Posters Pre/post treatment.

Data collection during my action research provoked more questions about how to support and scaffold students in argumentation. I am curious about the prompt/reasoning relationship and the role development plays when I look at the gains of sixth and seventh grade.

REFERENCES CITED

Bransford, J. D., Brown, A. I., & Cocking, R. R. (1999). *How people learn*. National Academy Press.

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional content. *Science Education*, 765-793.

Bohr, N. (1934). *Atomic theory and the description of nature.* Introduction.

Christensen, C. R., Garvin, D. A., & Sweet, A. (1991) *Education for judgment*. Harvard Business School Press.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64-74.

Enderle, P., Bickel, R., Gleim, L., Granger, E., Grooms, J., Hester, M., … Southerland, S. A. (2015). In *Argument-driven inquiry in life science: Lab investigations for grades 6-8*, 72–92. NSTA Press.

Kussmaul, C., & Sullivan, M. (2019). Activity selection and writing. In S. R. Simonson (Ed.), *POGIL: An introduction to process oriented guided inquiry learning for those who wish to empower learners* (pp.141-170). Stylus Publishing LLC.

Meacham, B. (2017). Implementing the claim, evidence, reasoning in the chemistry classroom. https://www.chemedx.org/article/implementing-claim-evidence-reasoning-framework-chemistry-chemistry-classroom.

McNeill, K. L. & Martin, D. M. (2011, April/May). Claim, evidence, and reasoning. *Science and Children*, 52-56.

Moog, R. S., Spencer J. N., & Straumanis, A. R. (2006). Process-oriented guided inquiry learning: POGIL and the POGIL project. *Metropolitan Universities*, 41-49.

Moog, R. S. (2019). Origins of POGIL: Process oriented guided inquiry learning. In S. R. Simonson (Ed.), *POGIL: An introduction to process oriented guided inquiry learning for those who wish to empower learners*, 23-41. Stylus Publishing LLC.

National Research Council. (1995). National Educational Standards. National Academy Press.

National Research Council. (2012). *A framework for K-12 science education: practices, cross-cutting concepts, and core ideas*. National Academy Press.

Sampson, V., Grooms J. R., Walker J. P. (2009, November). Argument-driven inquiry: A way to promote learning during laboratory activities. *The Science Teacher*, 42-47.

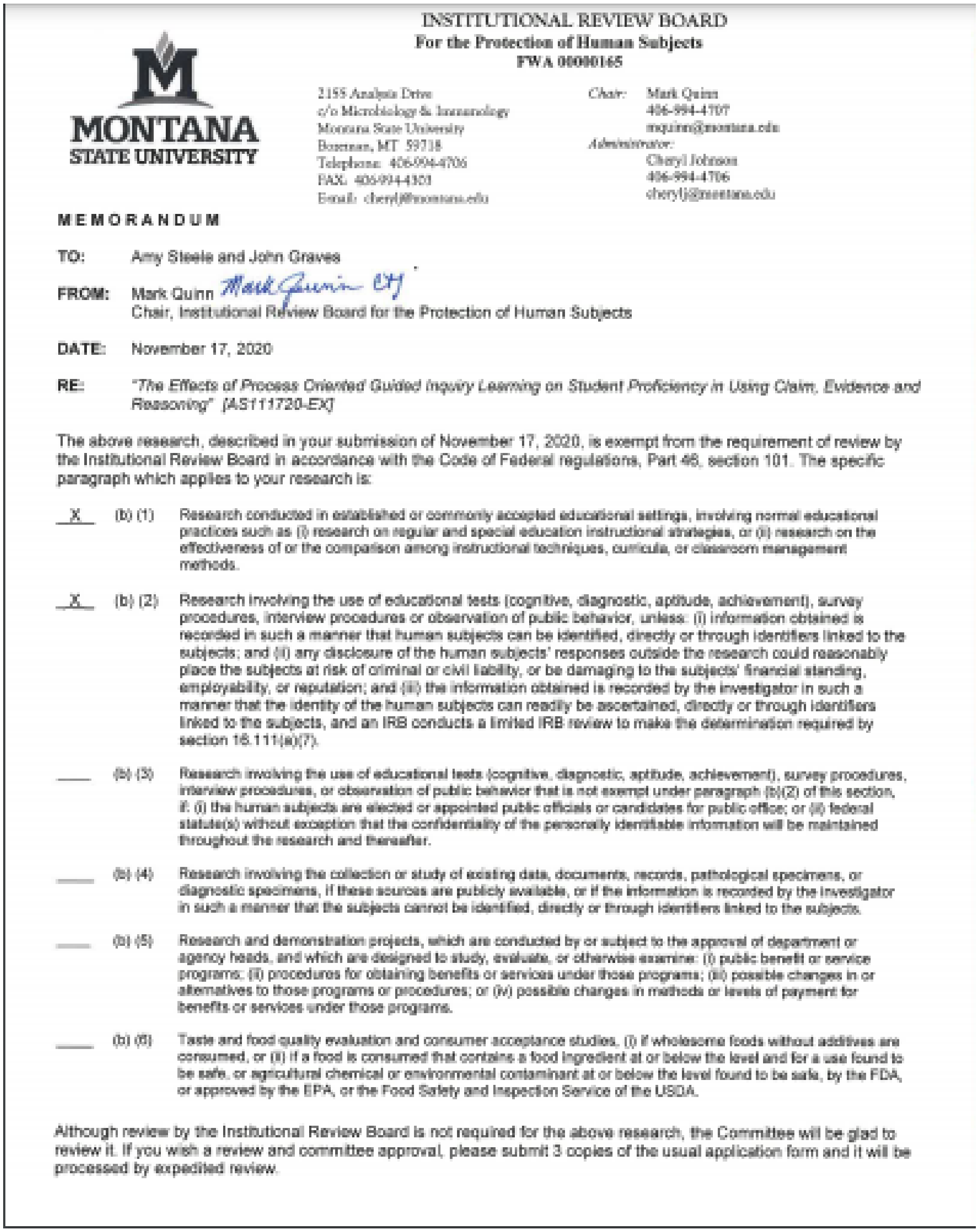
Sampson, V., Murphy, A., Lipscomb, K., & Hutner, T. (2018). In *Argument-driven inquiry in earth and space science: ab investigations for grades 6-10*, 310–330. NSTA Press.

Vellom, R. P. & Anderson, C. W. (1999). Reasoning about data in middle school science. *Journal of Research in Science Teaching*, 179-199.

APPENDICES

APPENDIX A

IRB EXEMPTION



INSTITUTIONAL REVIEW BOARD
For the Protection of Human Subjects
FWA 00000165

MONTANA
STATE UNIVERSITY

2155 Analysis Drive
c/o Microbiology & Immunology
Montana State University
Bozeman, MT 59718
Telephone: 406-994-4706
FAX: 406-994-4303
Email: cherylj@montana.edu

Chair: Mark Quinn
406-994-4707
mquinn@montana.edu
Administrator:
Cheryl Johnson
406-994-4706
cherylj@montana.edu

MEMORANDUM

TO: Amy Steele and John Graves

FROM: Mark Quinn Mark Quinn CJ
Chair, Institutional Review Board for the Protection of Human Subjects

DATE: November 17, 2020

RE: *"The Effects of Process Oriented Guided Inquiry Learning on Student Proficiency in Using Claim, Evidence and Reasoning" [AS111720-EX]*

The above research, described in your submission of November 17, 2020, is exempt from the requirement of review by the Institutional Review Board in accordance with the Code of Federal regulations, Part 46, section 101. The specific paragraph which applies to your research is:

X (b) (1) Research conducted in established or commonly accepted educational settings, involving normal educational practices such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

X (b) (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless: (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation; and (iii) the information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by section 16.111(a)(7).

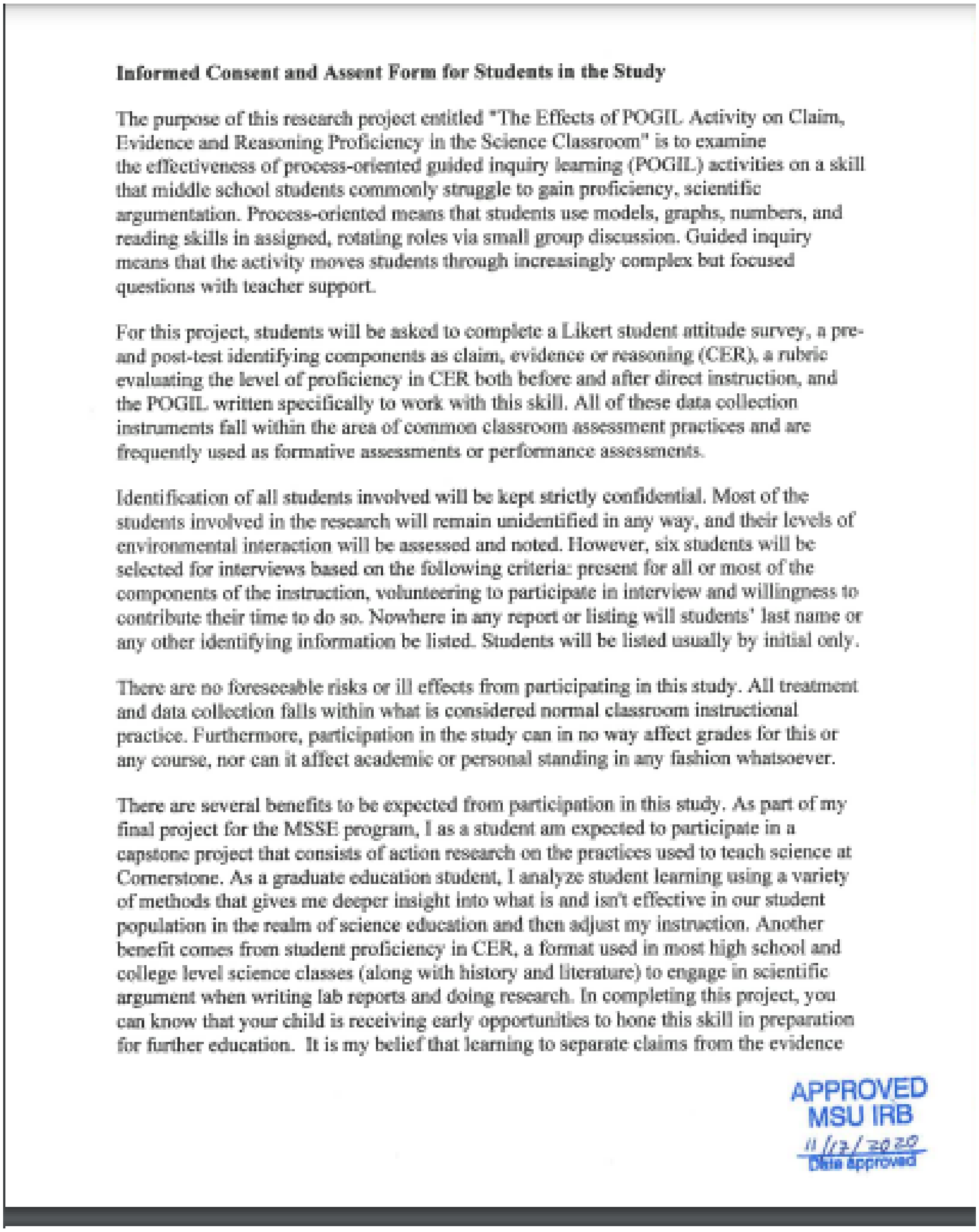
____ (b) (3) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under paragraph (b)(2) of this section, if: (i) the human subjects are elected or appointed public officials or candidates for public office; or (ii) federal statute(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.

____ (b) (4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available, or if the information is recorded by the investigator in such a manner that the subjects cannot be identified, directly or through identifiers linked to the subjects.

____ (b) (5) Research and demonstration projects, which are conducted by or subject to the approval of department or agency heads, and which are designed to study, evaluate, or otherwise examine: (i) public benefit or service programs; (ii) procedures for obtaining benefits or services under those programs; (iii) possible changes in or alternatives to those programs or procedures; or (iv) possible changes in methods or levels of payment for benefits or services under those programs.

____ (b) (6) Taste and food quality evaluation and consumer acceptance studies, (i) if wholesome foods without additives are consumed, or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the FDA, or approved by the EPA, or the Food Safety and Inspection Service of the USDA.

Although review by the Institutional Review Board is not required for the above research, the Committee will be glad to review it. If you wish a review and committee approval, please submit 3 copies of the usual application form and it will be processed by expedited review.



Informed Consent and Assent Form for Students in the Study

The purpose of this research project entitled "The Effects of POGIL Activity on Claim, Evidence and Reasoning Proficiency in the Science Classroom" is to examine the effectiveness of process-oriented guided inquiry learning (POGIL) activities on a skill that middle school students commonly struggle to gain proficiency, scientific argumentation. Process-oriented means that students use models, graphs, numbers, and reading skills in assigned, rotating roles via small group discussion. Guided inquiry means that the activity moves students through increasingly complex but focused questions with teacher support.

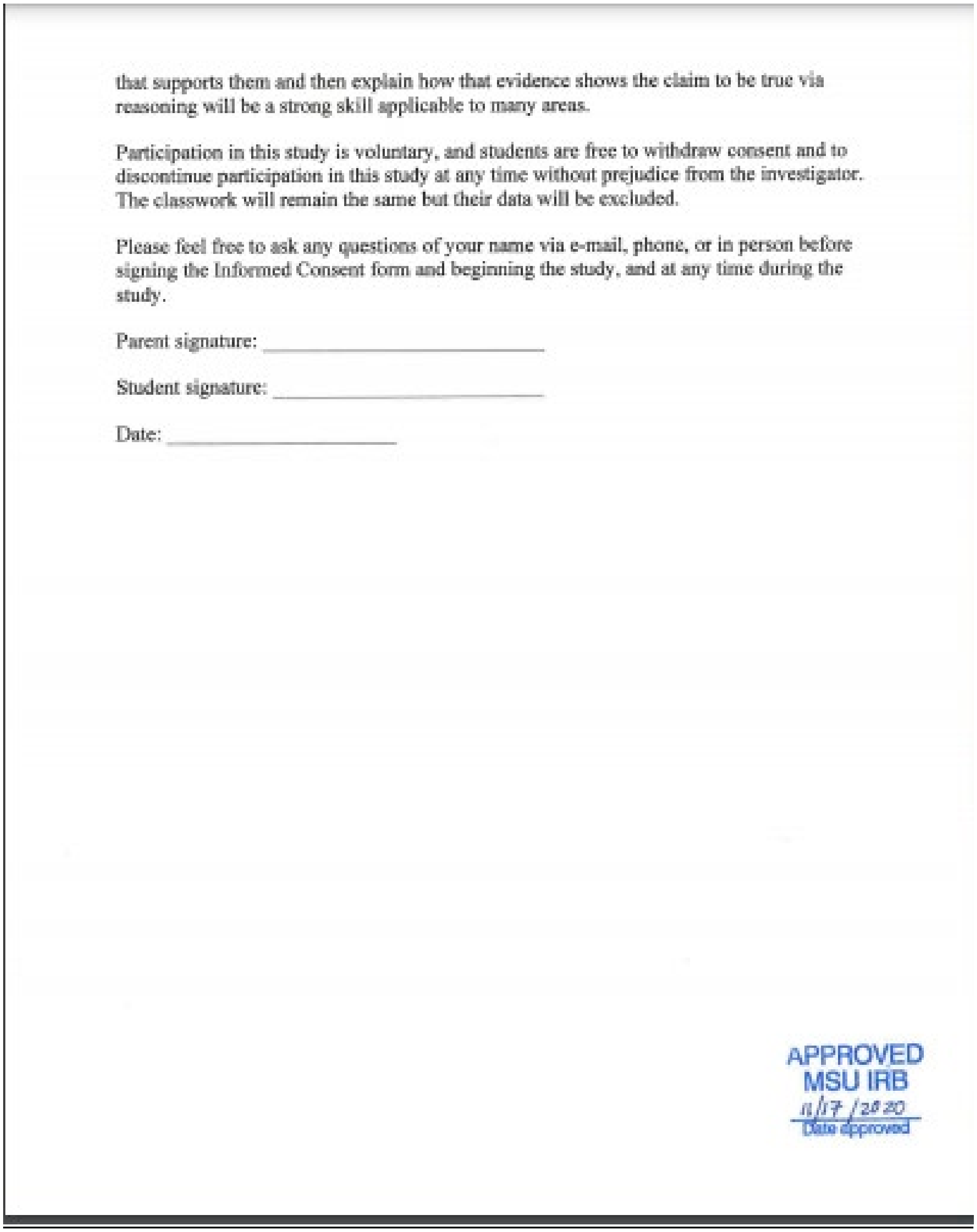
For this project, students will be asked to complete a Likert student attitude survey, a pre- and post-test identifying components as claim, evidence or reasoning (CER), a rubric evaluating the level of proficiency in CER both before and after direct instruction, and the POGIL written specifically to work with this skill. All of these data collection instruments fall within the area of common classroom assessment practices and are frequently used as formative assessments or performance assessments.

Identification of all students involved will be kept strictly confidential. Most of the students involved in the research will remain unidentified in any way, and their levels of environmental interaction will be assessed and noted. However, six students will be selected for interviews based on the following criteria: present for all or most of the components of the instruction, volunteering to participate in interview and willingness to contribute their time to do so. Nowhere in any report or listing will students' last name or any other identifying information be listed. Students will be listed usually by initial only.

There are no foreseeable risks or ill effects from participating in this study. All treatment and data collection falls within what is considered normal classroom instructional practice. Furthermore, participation in the study can in no way affect grades for this or any course, nor can it affect academic or personal standing in any fashion whatsoever.

There are several benefits to be expected from participation in this study. As part of my final project for the MSSE program, I as a student am expected to participate in a capstone project that consists of action research on the practices used to teach science at Cornerstone. As a graduate education student, I analyze student learning using a variety of methods that gives me deeper insight into what is and isn't effective in our student population in the realm of science education and then adjust my instruction. Another benefit comes from student proficiency in CER, a format used in most high school and college level science classes (along with history and literature) to engage in scientific argument when writing lab reports and doing research. In completing this project, you can know that your child is receiving early opportunities to hone this skill in preparation for further education. It is my belief that learning to separate claims from the evidence

APPROVED
MSU IRB
11/12/2020
Date approved



that supports them and then explain how that evidence shows the claim to be true via reasoning will be a strong skill applicable to many areas.

Participation in this study is voluntary, and students are free to withdraw consent and to discontinue participation in this study at any time without prejudice from the investigator. The classwork will remain the same but their data will be excluded.

Please feel free to ask any questions of your name via e-mail, phone, or in person before signing the Informed Consent form and beginning the study, and at any time during the study.

Parent signature: ___________________________

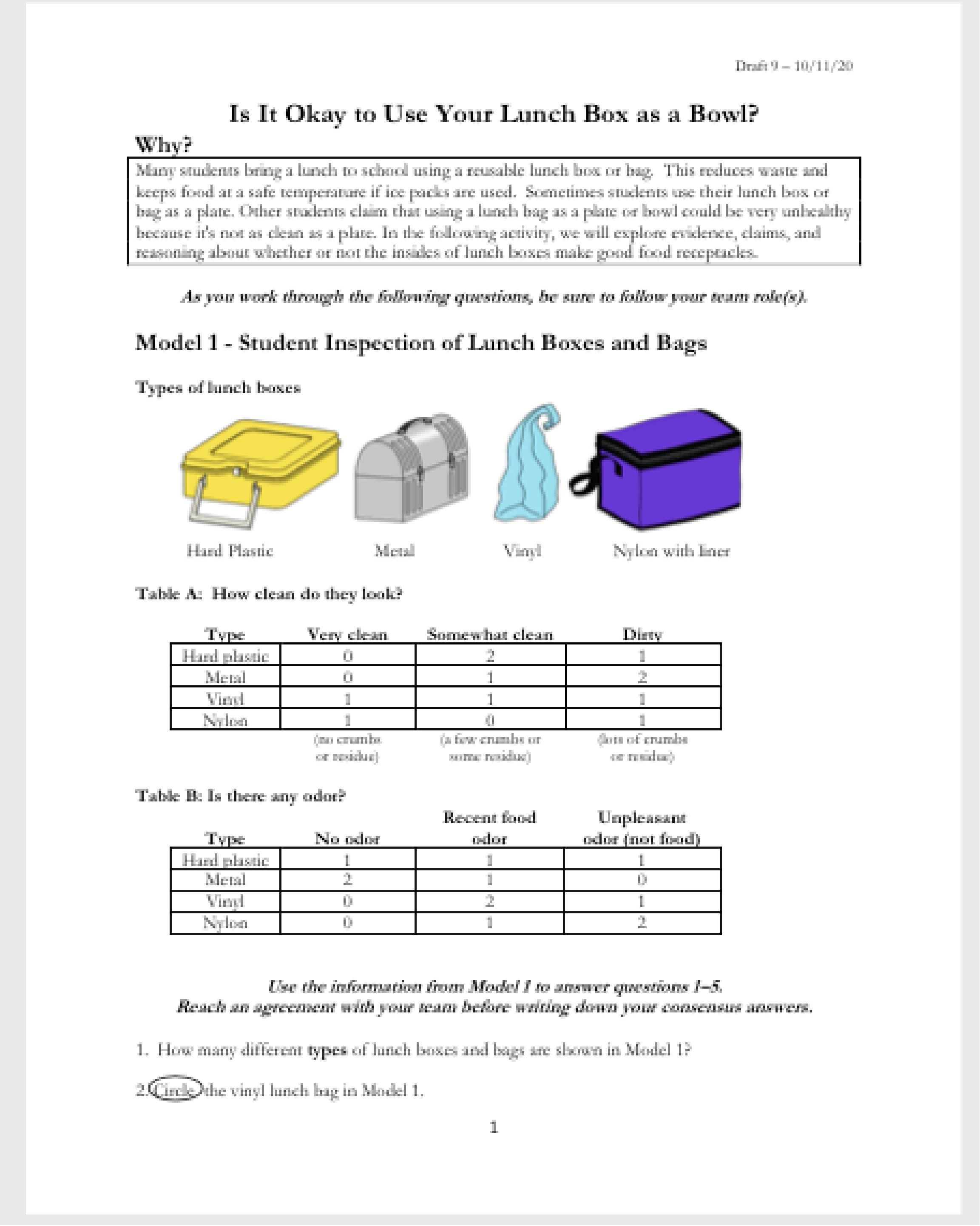
Student signature: __________________________

Date: ______________________

APPROVED
MSU IRB
11/17/2020
Date approved

APPENDIX B

CER POGIL



Draft 9 – 10/11/20

Is It Okay to Use Your Lunch Box as a Bowl?

Why?

Many students bring a lunch to school using a reusable lunch box or bag. This reduces waste and keeps food at a safe temperature if ice packs are used. Sometimes students use their lunch box or bag as a plate. Other students claim that using a lunch bag as a plate or bowl could be very unhealthy because it's not as clean as a plate. In the following activity, we will explore evidence, claims, and reasoning about whether or not the insides of lunch boxes make good food receptacles.

As you work through the following questions, be sure to follow your team role(s).

Model 1 - Student Inspection of Lunch Boxes and Bags

Types of lunch boxes

Hard Plastic Metal Vinyl Nylon with liner

Table A: How clean do they look?

Type	Very clean	Somewhat clean	Dirty
Hard plastic	0	2	1
Metal	0	1	2
Vinyl	1	1	1
Nylon	1	0	1
	(no crumbs or residue)	(a few crumbs or some residue)	(lots of crumbs or residue)

Table B: Is there any odor?

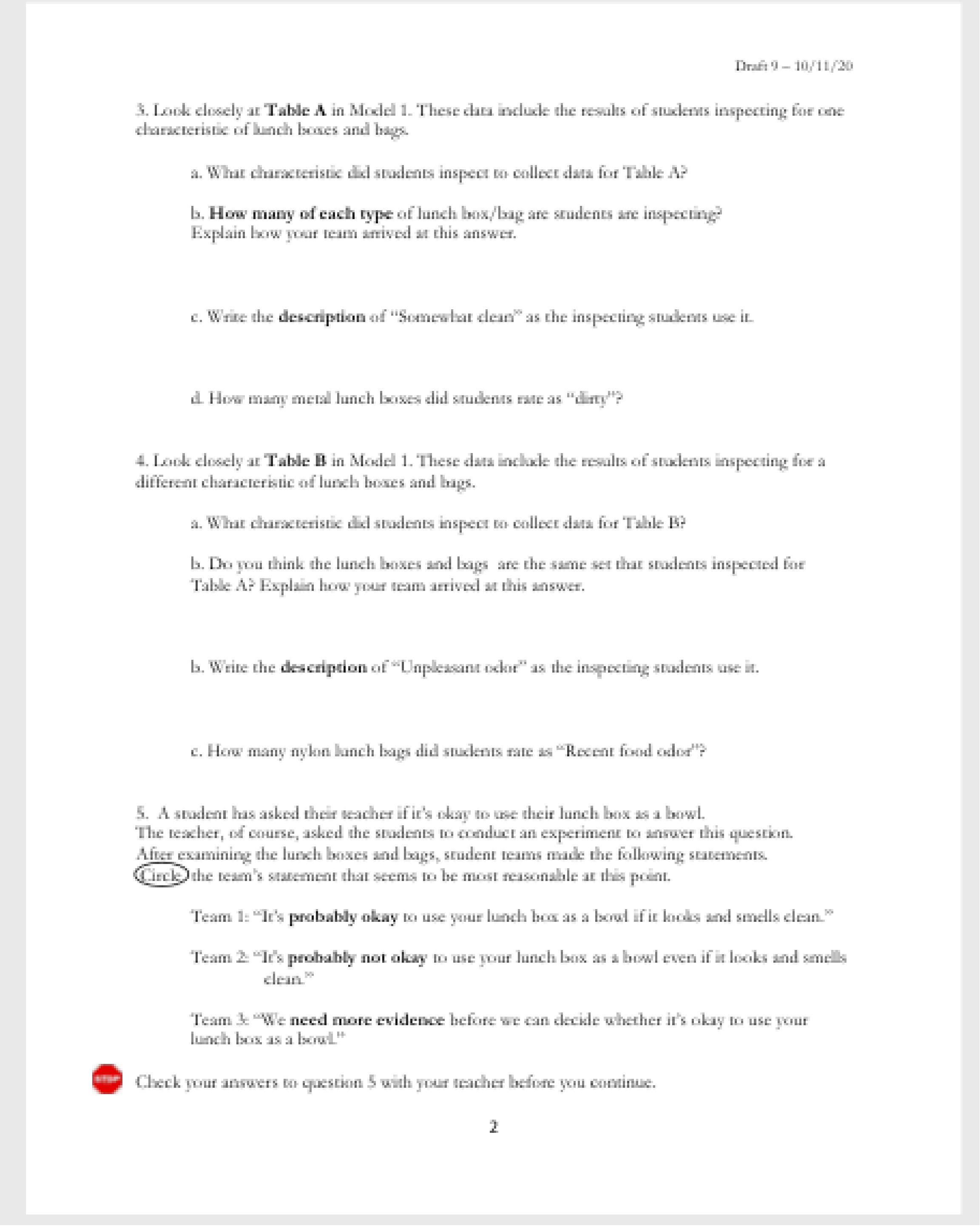
Type	No odor	Recent food odor	Unpleasant odor (not food)
Hard plastic	1	1	1
Metal	2	1	0
Vinyl	0	2	1
Nylon	0	1	2

Use the information from Model 1 to answer questions 1–5.
Reach an agreement with your team before writing down your consensus answers.

1. How many different **types** of lunch boxes and bags are shown in Model 1?

2. Circle the vinyl lunch bag in Model 1.

1



3. Look closely at **Table A** in Model 1. These data include the results of students inspecting for one characteristic of lunch boxes and bags.

a. What characteristic did students inspect to collect data for Table A?

b. **How many of each type** of lunch box/bag are students are inspecting?
Explain how your team arrived at this answer.

c. Write the **description** of "Somewhat clean" as the inspecting students use it.

d. How many metal lunch boxes did students rate as "dirty"?

4. Look closely at **Table B** in Model 1. These data include the results of students inspecting for a different characteristic of lunch boxes and bags.

a. What characteristic did students inspect to collect data for Table B?

b. Do you think the lunch boxes and bags are the same set that students inspected for Table A? Explain how your team arrived at this answer.

b. Write the **description** of "Unpleasant odor" as the inspecting students use it.

c. How many nylon lunch bags did students rate as "Recent food odor"?

5. A student has asked their teacher if it's okay to use their lunch box as a bowl.
The teacher, of course, asked the students to conduct an experiment to answer this question.
After examining the lunch boxes and bags, student teams made the following statements.
Circle the team's statement that seems to be most reasonable at this point.

Team 1: "It's **probably okay** to use your lunch box as a bowl if it looks and smells clean."

Team 2: "It's **probably not okay** to use your lunch box as a bowl even if it looks and smells clean."

Team 3: "We **need more evidence** before we can decide whether it's okay to use your lunch box as a bowl."

Check your answers to question 5 with your teacher before you continue.

Read This!

Most bacteria are not dangerous to humans. However, some bacteria might make you sick if you accidentally eat them. Scientists look for these "bad" bacteria on surfaces by wiping the surface with a cotton swab. Then they wipe the swab on a small dish of bacteria food. After the bacteria grow for about two days, the scientists look for small pinkish-purple dots that represent an entire colony of many millions of bacteria. If someone eats more than 10 colonies, they can get sick.

Model 2 - Bacteria Found in Lunch Boxes and Bags

Diagram: Collecting and analyzing samples to detect presence of bacteria

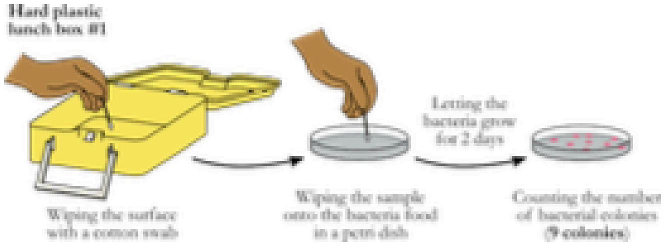
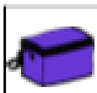
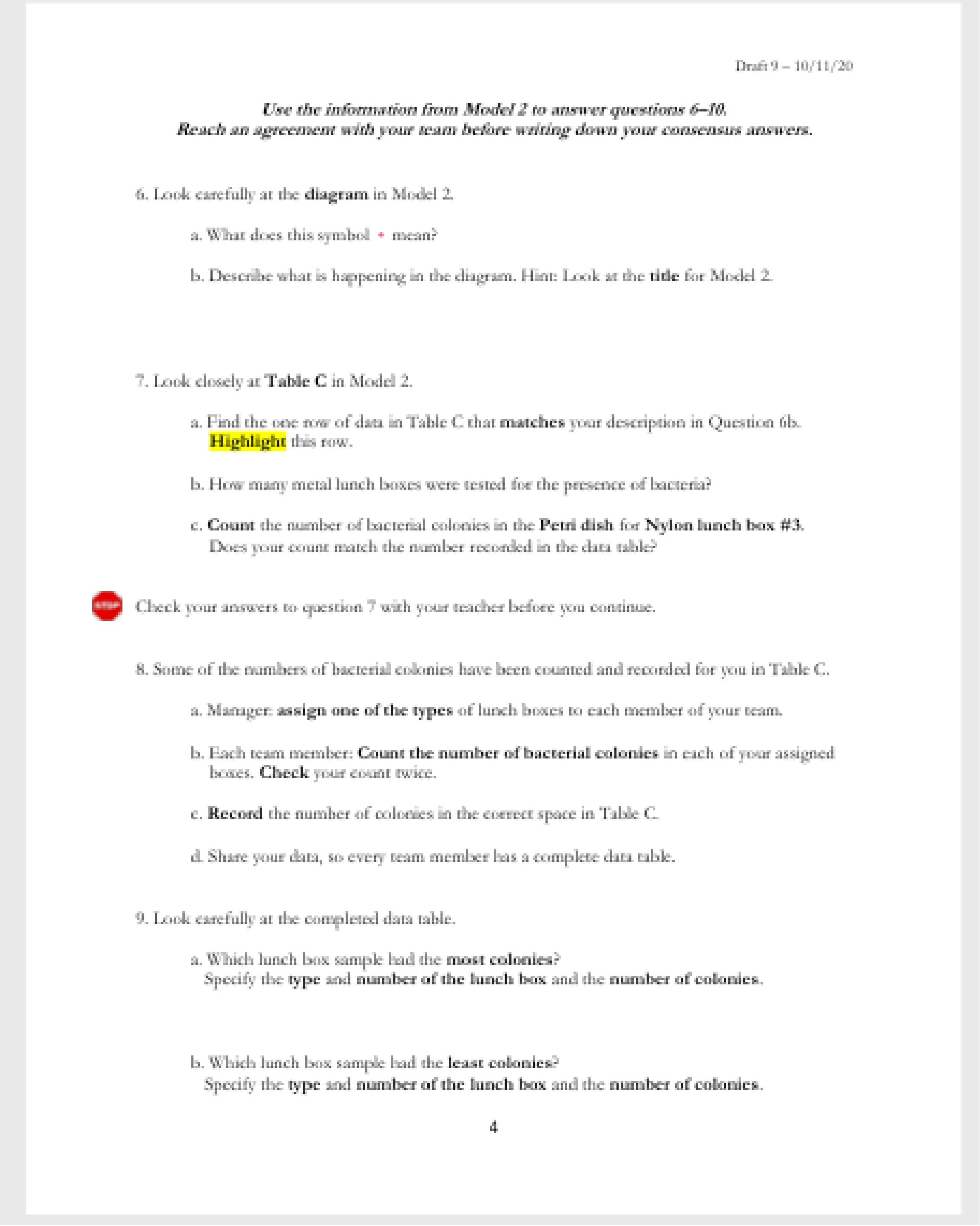


Table C: Number of colonies found in different types of lunch boxes

Type	Box #	Petri dish	# of colonies	Type	Box #	Petri dish	# of colonies
 Hard Plastic	1		*9*	 Vinyl	1		*3*
	2				2		
	3				3		
 Metal	1			 Nylon	1		
	2		*13*		2		
	3				3		*16*



Use the information from Model 2 to answer questions 6–10.
Reach an agreement with your team before writing down your consensus answers.

6. Look carefully at the **diagram** in Model 2.

   a. What does this symbol * mean?

   b. Describe what is happening in the diagram. Hint: Look at the **title** for Model 2.

7. Look closely at **Table C** in Model 2.

   a. Find the one row of data in Table C that **matches** your description in Question 6b. **Highlight** this row.

   b. How many metal lunch boxes were tested for the presence of bacteria?

   c. **Count** the number of bacterial colonies in the **Petri dish** for **Nylon lunch box #3**. Does your count match the number recorded in the data table?

🛑 Check your answers to question 7 with your teacher before you continue.

8. Some of the numbers of bacterial colonies have been counted and recorded for you in Table C.

   a. Manager: **assign one of the types** of lunch boxes to each member of your team.

   b. Each team member: **Count the number of bacterial colonies** in each of your assigned boxes. **Check** your count twice.

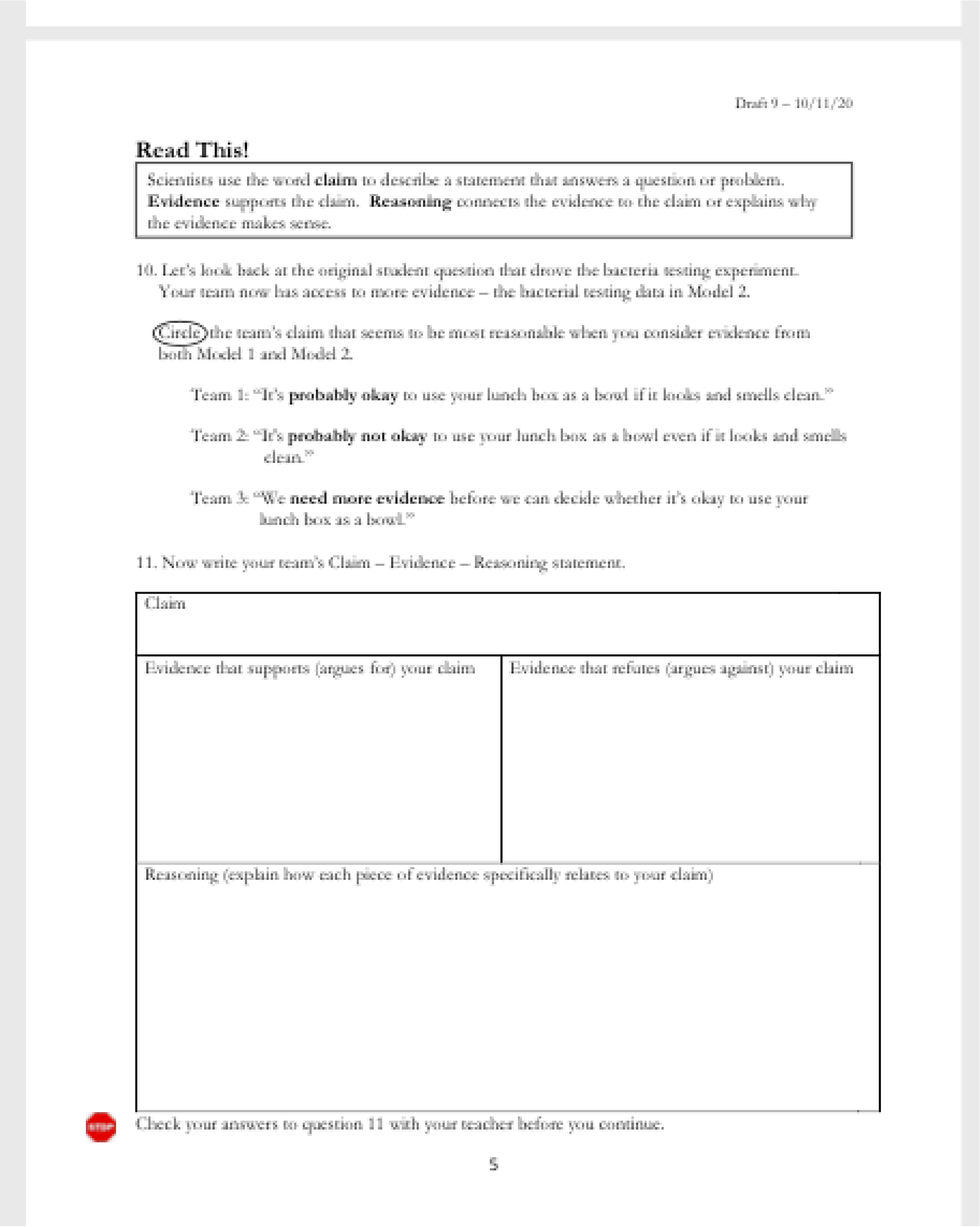
   c. **Record** the number of colonies in the correct space in Table C.

   d. Share your data, so every team member has a complete data table.

9. Look carefully at the completed data table.

   a. Which lunch box sample had the **most colonies**?
   Specify the **type** and **number of the lunch box** and the **number of colonies**.

   b. Which lunch box sample had the **least colonies**?
   Specify the **type** and **number of the lunch box** and the **number of colonies**.



Draft 9 – 10/11/20

Read This!

Scientists use the word **claim** to describe a statement that answers a question or problem. **Evidence** supports the claim. **Reasoning** connects the evidence to the claim or explains why the evidence makes sense.

10. Let's look back at the original student question that drove the bacteria testing experiment. Your team now has access to more evidence – the bacterial testing data in Model 2.

Circle the team's claim that seems to be most reasonable when you consider evidence from both Model 1 and Model 2.

Team 1: "It's **probably okay** to use your lunch box as a bowl if it looks and smells clean."

Team 2: "It's **probably not okay** to use your lunch box as a bowl even if it looks and smells clean."

Team 3: "We **need more evidence** before we can decide whether it's okay to use your lunch box as a bowl."

11. Now write your team's Claim – Evidence – Reasoning statement.

Claim	
Evidence that supports (argues for) your claim	Evidence that refutes (argues against) your claim
Reasoning (explain how each piece of evidence specifically relates to your claim)	

STOP Check your answers to question 11 with your teacher before you continue.

5

APPENDIX C

BACKGROUND KNOWLEDGE PROBE FOR CLAIM, EVIDENCE AND REASONING

Claim, Evidence and Reasoning

In science, we use the words **claim, evidence** and **reasoning** to describe parts of a scientific argument. What do you know about these three words?

APPENDIX D

STUDENT ATTITUDES REGARDING CER

Likert Survey for Claim, Evidence and Reasoning

DIRECTIONS: For each of the following statements, please choose the response that you feel applies to you best by coloring in once circle in each row.

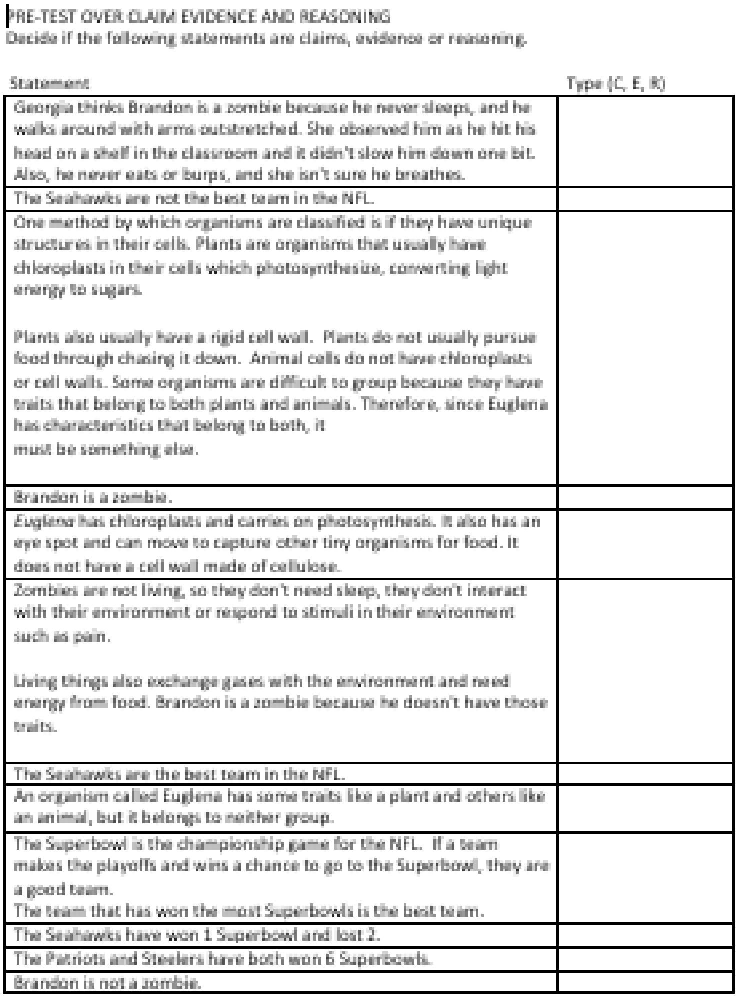
Question 7 gives you a chance to let me know anything else you feel I need to understand about your thinking and understanding about claim, evidence, and reasoning.

	Strongly agree	Agree	Disagree	Disagree Strongly
1. I can make a claim to answer a scientific question or explain a phenomenon.	○	○	○	○
2. I can identify a claim about a scientific phenomenon or question.	○	○	○	○
3. I can identify evidence used to support a scientific claim.	○	○	○	○
4. I can tell the difference between evidence and the reasoning that connects evidence to a claim.	○	○	○	○
5. I can assess evidence and determine what doesn't support a claim.	○	○	○	○
6. I can use reasoning to explain why a claim is supported by evidence.	○	○	○	○
7.What else would you like for me to know?				

APPENDIX E

PRE- AND POST-TEST:

IDENTIFYING CLAIM, EVIDENCE AND REASONING STATEMENTS

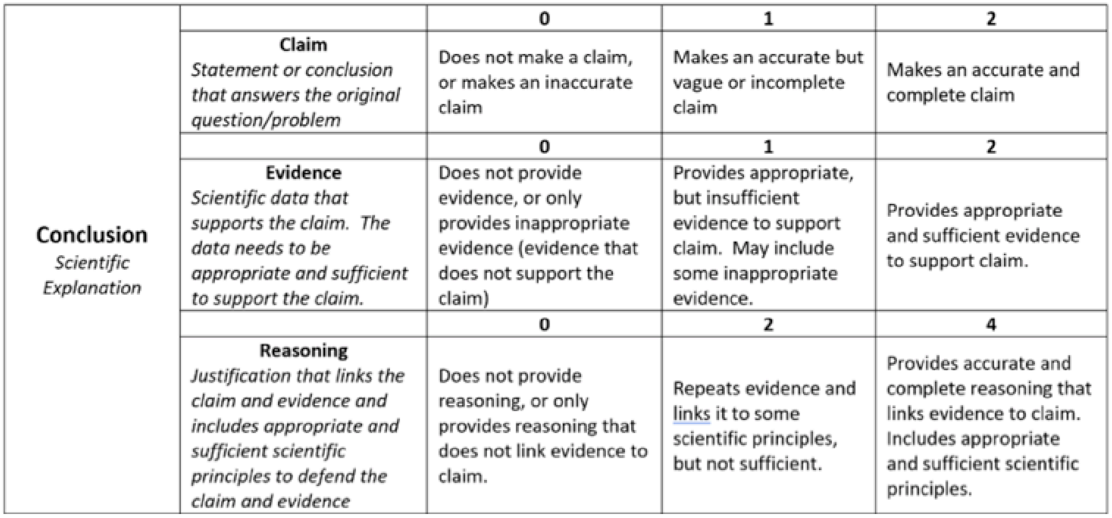


PRE-TEST OVER CLAIM EVIDENCE AND REASONING
Decide if the following statements are claims, evidence or reasoning.

Statement	Type (C, E, R)
Georgia thinks Brandon is a zombie because he never sleeps, and he walks around with arms outstretched. She observed him as he hit his head on a shelf in the classroom and it didn't slow him down one bit. Also, he never eats or burps, and she isn't sure he breathes.	
The Seahawks are not the best team in the NFL.	
One method by which organisms are classified is if they have unique structures in their cells. Plants are organisms that usually have chloroplasts in their cells which photosynthesize, converting light energy to sugars. Plants also usually have a rigid cell wall. Plants do not usually pursue food through chasing it down. Animal cells do not have chloroplasts or cell walls. Some organisms are difficult to group because they have traits that belong to both plants and animals. Therefore, since Euglena has characteristics that belong to both, it must be something else.	
Brandon is a zombie.	
Euglena has chloroplasts and carries on photosynthesis. It also has an eye spot and can move to capture other tiny organisms for food. It does not have a cell wall made of cellulose.	
Zombies are not living, so they don't need sleep, they don't interact with their environment or respond to stimuli in their environment such as pain. Living things also exchange gases with the environment and need energy from food. Brandon is a zombie because he doesn't have those traits.	
The Seahawks are the best team in the NFL.	
An organism called Euglena has some traits like a plant and others like an animal, but it belongs to neither group.	
The Superbowl is the championship game for the NFL. If a team makes the playoffs and wins a chance to go to the Superbowl, they are a good team. The team that has won the most Superbowls is the best team.	
The Seahawks have won 1 Superbowl and lost 2.	
The Patriots and Steelers have both won 6 Superbowls.	
Brandon is not a zombie.	

APPENDIX F

CER ASSESSMENT RUBRIC



Conclusion *Scientific Explanation*		**0**	**1**	**2**
	Claim *Statement or conclusion that answers the original question/problem*	Does not make a claim, or makes an inaccurate claim	Makes an accurate but vague or incomplete claim	Makes an accurate and complete claim
		0	**1**	**2**
	Evidence *Scientific data that supports the claim. The data needs to be appropriate and sufficient to support the claim.*	Does not provide evidence, or only provides inappropriate evidence (evidence that does not support the claim)	Provides appropriate, but insufficient evidence to support claim. May include some inappropriate evidence.	Provides appropriate and sufficient evidence to support claim.
		0	**2**	**4**
	Reasoning *Justification that links the claim and evidence and includes appropriate and sufficient scientific principles to defend the claim and evidence*	Does not provide reasoning, or only provides reasoning that does not link evidence to claim.	Repeats evidence and links it to some scientific principles, but not sufficient.	Provides accurate and complete reasoning that links evidence to claim. Includes appropriate and sufficient scientific principles.