

CLAIM, EVIDENCE, REASONING IN MIDDLE SCHOOL SCIENCE

by

Miriam Gail Harms

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

of

Master of Science

in

Science Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2018

©COPYRIGHT

by

Miriam Gail Harms

2018

All Rights Reserved

TABLE OF CONTENTS

1. INTRODUCTION AND BACKGROUND1

2. CONCEPTUAL FRAMEWORK5

3. METHODOLOGY9

4. DATA AND ANALYSIS18

5. INTERPRETATION AND CONCLUSION36

6. VALUE39

REFERENCES CITED42

APPENDICES

 APPENDIX A IRB Exemption46

 APPENDIX B Ecology and the Environment Test48

 APPENDIX C After-Test Survey56

 APPENDIX D Is air matter?58

 APPENDIX E CER Coding60

 APPENDIX F Post Interview Questions64

LIST OF TABLES

1. Eighth-grade CER Lab Activities	13
2. Outcome Space for How does mass affect friction?	14
3. Seventh-grade CER Lab Activities	17
4. Data Triangulation Matrix	18

LIST OF FIGURES

1. The Core Four Elements of Personalized Learning3

2. Claim, Evidence, Reasoning Example.....8

3. CER Practice about Cockroaches16

4. Comparative Boxplots for Eighth-grade Test Scores19

5. Average Student Confidence from the After-Test Survey.....20

6. Eighth-grade Coding for Claims Made on Labs21

7. Eighth-grade Coding for Evidence on Labs.....22

8. Eighth-grade Coding for Reasoning on Labs.....23

9. Eighth-grade Coding for Claims on Tests24

10. Eighth-grade Coding for Evidence on Tests24

11. Eighth-grade Coding for Reasoning on Tests.....25

12. Comparative boxplots for Seventh-grade Test Scores.....26

13. Seventh-grade Coding for Claims on Labs27

14. Seventh-grade Coding for Evidence on Labs29

15. Seventh-grade Coding for Reasoning on Labs30

16. Seventh-grade Coding for Claims on Assessments31

17. Seventh-grade Coding for Evidence on Assessments.....32

18. Seventh-grade Coding for Reasoning on Assessments.....33

19. Student A’s CER Response on the Friction Lab.....35

20. Student A’s CER Response on the Magnetism Lab35

21. Student A’s CER Response on Is Air Matter?.....36

ABSTRACT

A technique for writing evidence-based claims called claim, evidence, and reasoning was integrated into a middle school science curriculum with the purpose of improving students' test scores. This study also sought to see the effect of this technique on students' testing confidence and success with writing a claim, evidence, and reasoning. Through a pre- and post-treatment part of the study, student test scores were collected and students completed a post-test survey to gauge students' confidence. The claim, evidence, and reasoning written by students at the end of labs were coded to track the quality of student responses. The answers that students gave to short answer test questions were also coded to assess their integration of claim, evidence, and reasoning into their test answers. Students' test scores and testing confidence did not improve, but students increased their use of the claim, evidence, and reasoning skill on tests. Students also improved the quality of their responses on labs where they were prompted to write a claim, evidence, and reasoning.

INTRODUCTION AND BACKGROUND

Before I became a teacher, I considered the claim made by students that they were “just not good at tests” to be a lazy excuse. When I was in high school and college, fellow students would utter this phrase to explain away their poor test grades. I never struggled with taking tests and felt that they should study more instead of making excuses. It seemed to me that students used their frustration with tests as a reason to try less. After the time I’ve spent as a middle and high school science teacher, I recognize while some students do poorly on tests because they need more study time, this is not always the case, and some students simply are bad at taking tests. I have students whom I am confident know the material but I am repeatedly baffled at how poorly they perform on exams. They are unable to show what they know on a test consisting of multiple choice and short answer questions.

My initial response to this question was to vary the ways I assess students to better determine what students actually understand. However, I have an obligation to prepare my students for a science standardized test they take at the end of eighth-grade. The year before I began my capstone project, I began to spend more time on subject related vocabulary and high-frequency vocabulary where the meanings in science are different than in everyday language. As I worked to build vocabulary with my students, I also instituted something called Friday Skills in my seventh-grade physical science classes during the 2016-2017 school year and continued it this school year. The majority of the eighth-grade students in my study participated in Friday Skills during their seventh-grade year and I have improved the practice with my current seventh-grade

students. Each Friday, we spend the first fifteen minutes of class practicing an academic or science skill. These cover a wide range of topics including reading graphs and tables, determining the relationship between variables, identifying distractors in multiple choice questions, and making observations and inferences.

Although I saw marginal improvement in my students' academic performance last year with my increased focus on vocabulary and science skills, there was progress to be made. Over my three years of teaching, I have read many short answer responses in which students do not do much more than restate the question. This is mirrored in my classroom where it often takes several follow up questions of, but "why?" to get students to dig into what they know and think of the reasoning behind the claim they are making. I have found my students are not programmed to support their claims with reasoning and I believe this is a major factor in them not thinking through all test questions, not just short answers questions. When I sit down with students to go through test questions they have missed, they can easily answer questions correctly when I prompt them to explain to me the reasoning for their answers. I often find that in their explanation of an incorrect answer they begin to realize why they were wrong and what the correct answer was. However, thinking through their reasoning is not something my students are used to doing.

I teach in a small rural school district in Romulus, New York. Around 400 students attend our district's public pre-K to 12 school and our students benefit from small classes (U.S. News and World Report, 2014). Our district is unique with a 1:1 chrome-book program and a Personalized Learning initiative. Our school has contracted

a company named Education Elements and our administration is supporting us in a transition to a pedagogy that involves the needs of each learner by integrating Education Element's Core Four of targeted instruction, data-driven decisions, integrated digital content, and student reflection and ownership (Johns & Wolking, 2018) (Figure 1). These changes have interfered with my research but without complaint from me. The Personalized Learning related changes I am making in my classroom are drastically improving my students' independence and ownership in their work. In turn, this improves their learning.

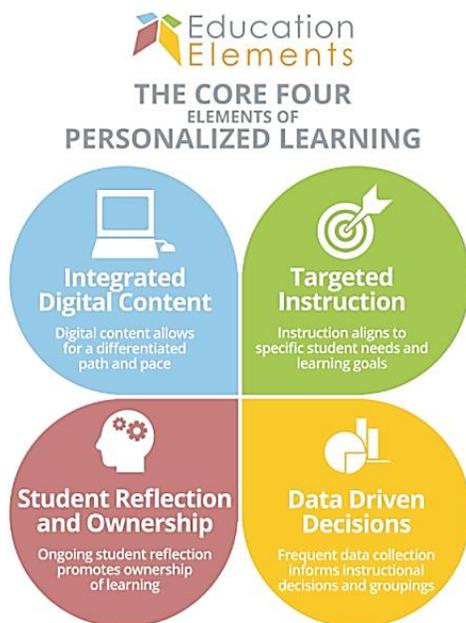


Figure 1. The Core Four Elements of Personalized Learning (Johns & Wolking, 2018).

In New York, the state education department is rolling out a new set of science standards based on the Next Generation Science Standards (NGSS Lead States, 2013). One of the new aspects of science education that NGSS is bringing to my classroom is the focus on science and engineering practices. One of these practices is argumentation. In this practice, students are expected to make scientific claims and support them with

evidence and reasoning. Students also use this skill to defend their claims while participating in scientific argumentation in the classroom.

Before the capstone project when I began to implement the practice of argumentation into my classes in a few small ways, I was impressed with my students' ability to work together to develop the best evidence and reasoning for their claims. Anecdotally, I saw students' understanding of topics solidify as they were stretched to show how they knew something and then explain why that evidence supported their claim. It is for that reason that I decided to increasingly implement argumentation in my classroom because I believe that the practice of argumentation will also benefit my students' ability to take multiple choice and short answer tests. With continued use of the claim, evidence, reasoning (CER) process, I hoped to make the use of these higher level thinking skills more automatic for my students.

With the CER framework, students begin by answering a question with a claim. Students follow the claim with the evidence that supports it. In most cases, this evidence is data from an experiment they have conducted or observations they have made. The most difficult step for students, but also the most integral, is the reasoning step. In their reasoning, students must connect the evidence to their claim by explaining how the evidence supports their claim and include scientific principles that make their claim true. For example, in a lesson on density, I had students complete a CER with sentence starters for several density questions. The students had completed an online simulation to help them investigate this question and presented their results in the form of CER. One of the questions was "how can you tell if a crown is made of pure gold?" After working on the

simulation and completing a CER organizer one group responded with the following: we know a crown is made of pure gold if it has the same density as gold. Our evidence is that the density of gold is 19.3 g/mL and the crowns that were not pure gold did not have that same density. We know this is true because density is a physical property of a substance and does not change, so the density of gold is always 19.3 g/mL.

My interest in the ability of CER to improve my students' test scores led to the creation of my focus statement: Will purposeful teaching of CER skills and repeated use of CER in inquiry experiments and assessments improve student test scores on traditional tests?

1. Will it improve student confidence on tests?
2. Will it improve student success with the CER technique?

CONCEPTUAL FRAMEWORK

Engaging students in inquiry is integral to teaching science. Inquiry models the work scientists do and, with an inquiry-based approach, students are invited to join in on the conversation. Student-led inquiry is not the same as doing experiments in the classroom. Students must learn to organize what they already know, identify what they'd like to know, generate a model, seek evidence, and construct an argument. In doing these things, students develop a deeper understanding of the content and make better connections with the content they are learning (Windschitl, 2008).

Learning about science cannot be about just knowing facts and content. It is necessary for students to develop scientific practices and ways of thinking about science. These skills are not easily picked up by students and need to be intentionally taught.

These practices should be as much of a priority and a goal as content learning (McNeill, 2011).

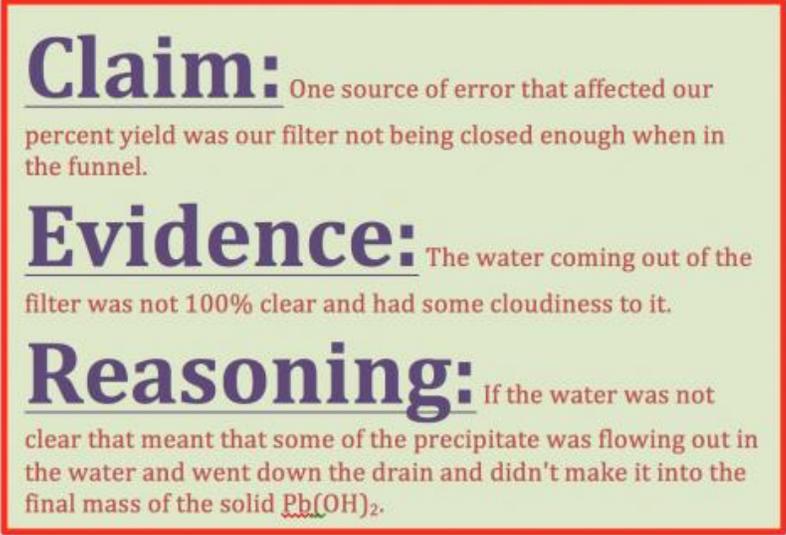
An important step in the inquiry process is having students participate in the construction of explanations and arguments. During the process of argumentation, students work together to explain phenomenon and evidence (McNeill, 2011). Studies confirm the need to teach argumentation. One such study began by training teachers in methods of teaching argumentation and providing evidence. This study was based on previous evidence that showed adolescents do not have adequate argumentation and higher-level thinking skills. The study showed that the students in classes where teachers used these new teaching techniques significantly improved their argumentation skills, over students whose teachers did not use these techniques. This evidence supports the need to teach argumentation as part of science curriculum (Bulgren, Ellis, & Marquis, 2014).

Educators who are preparing their students for content-focused standardized testing often have a concern about taking time away from content instruction. Educational research performed in China investigated whether there was a correlation between argumentation techniques and performance on mandatory rigorous high school testing. Their goal was to collect evidence to support the instructor's desire to spend more time teaching science practices in the classroom. Classes in the study received instruction in argumentation techniques and practiced them. The researchers found that taking the time to focus on improving argumentation skills did not improve or hinder standardized testing scores. Although improving argumentation skills did not help

student performance on testing, taking time away from content to focus on practices did not limit the students' performance. Additionally, the study found that learning argumentation skills were most beneficial to students with average test scores. These students showed significant improvement in written and verbal argumentation skills when compared with students with high test scores (Wang & Buck, 2015).

Argumentation and explanation go hand in hand. When students participate in argumentation, they learn to provide better explanations through improving their claims by justifying them through evidence and reasoning. The role that argumentation has in improving explanatory skills is that students learn through critiquing other students' explanations and use of evidence. In a study of elementary school students, it was found that through practice and explicit instruction, the science students were more likely to provide evidence and reasoning to support their claim without being prompted to do so (McNeill, 2011).

One technique used to support evidence and argumentation is called claim, evidence, reasoning or CER. This framework helps students identify their claim and support it with evidence (Figure 2). The evidence can be qualitative or quantitative and is used to decide on a claim and to support it. The reasoning is used to connect evidence to a claim through the scientific principles the students know (McNeill, 2011).



Claim: One source of error that affected our percent yield was our filter not being closed enough when in the funnel.

Evidence: The water coming out of the filter was not 100% clear and had some cloudiness to it.

Reasoning: If the water was not clear that meant that some of the precipitate was flowing out in the water and went down the drain and didn't make it into the final mass of the solid Pb(OH)_2 .

Figure 2. Claim, Evidence, Reasoning example (Meacham, 2017).

One study on the effectiveness of CER followed a fifth-grade science class as they used inquiry and argumentation over the course of a year. The teacher used varying methods to support the use of CER and to help students build argumentation skills. As they continued to practice constructing explanations and participating in argumentation, students became more sophisticated in the way they challenged each other's arguments (Chen, Hand, & Park, 2016). The students began to focus on the "coherence of the argument and the quality of evidence" (Chen, Hand, & Park, 2016, p. 296). It is clear that CER along with group and whole class argumentation helped these students understand how to provide quality evidence and how to connect that evidence to a claim (Chen, Hand, & Park, 2016). This research supports the effectiveness of CER to teach students both evidence and argumentation skills.

A great deal of research has been done as to determine the best ways to teach argumentation and the CER skill. One study shows that students need explicit instruction and guidelines for both argumentation and CER. Students also need time to internalize

and reflect on the content as well as time to improve their skills (Aydeniz & Bilican, 2016). Another researcher studied the effect that classroom practices have on argumentation. They concluded that a competitive environment is not best for argumentation. Reasoning skills and argumentation require students to work collaboratively and individually. The researcher looked for ways that the students' argumentation abilities were affected by classroom practices. Students should be taught and encouraged to justify their responses both verbally and in written arguments using claim, evidence, and reasoning. Students who were able to justify their claims were better engaged in argumentation in the science classroom. When students were encouraged to be competitive, they tried to be the student with the correct answer instead of working as a group to find the best explanations and arguments (Berland, 2011).

METHODOLOGY

This study integrated argumentation skills and the claim, evidence, reasoning (CER) framework into the curriculum of my three middle school science classes as a strategy to improve test scores. Additionally, this project was designed to see if the intentional teaching of CER skills would improve student confidence on tests and success with the CER technique.

The treatment group consisted of 15 eighth-grade life science students and 32 seventh-grade physical science students. The seventh-grade students were divided into two classes, a regular class and an advanced class. Both classes learned the same material, but at a different pace that allows for enrichment activities in the advanced class. I did not have two groups of eighth-grade students because our school advances a

group of students to start high school biology in eighth-grade. Additionally, I have included several students in this study who were not present during the whole school year. The research methodology for this project received an exemption by Montana State University's Institutional Review Board and compliance for working with human subjects was maintained (Appendix A).

Twelve of the eighth-grade students from this study were my students for their seventh-grade year as well. During their seventh-grade year, they had some exposure to the CER framework but it was not introduced systematically or extensively. Their English and social studies classes further acquainted them with the framework. The other teachers on our middle-level team wanted to unify the language we used with our students and throughout the school year used the CER framework when teaching students to use evidence-based claims. The treatment portion of the study is vastly different for the eighth-grade and seventh-grade groups. The seventh-graders were added later in the process after it became clear that my work with personalized learning was positively impacting the eighth-grade class and would make the influences of CER difficult to discern.

Eighth-grade Methodology

For my eighth-grade group, the year began with one unit of pre-treatment where CER was not introduced or used. This unit, Ecology and the Environment, concluded with a unit test (Appendix B). During each of the units throughout the pre-treatment and treatment portion of this study, student unit test scores were collected. All questions on the assessments for both seventh and eighth-grade were taken from past New York State

Intermediate Science Tests that are administered at the end of the eighth-grade year and cover the science material from fifth through eighth-grades. Eighth-grade students also completed a Post-Test Survey to assess their confidence on various aspects of the test and repeated this survey after each test (Appendix C). The Likert questions from the survey were reported using histograms that show the percentage of students with different levels of confidence throughout the study. The Likert questions were scored on a scale with a 5 being assigned to “strongly agree,” 4 for “agree,” 3 for “neutral,” 2 for “disagree,” and 1 for “strongly disagree.”

The treatment portion with the eighth-grade began with introduction to CER and students practiced the skill using examples that were familiar to them. Our first activity was to analyze the question “is air matter?” and students conducted a simple experiment by massing a balloon before and after it was blown up (Appendix D). This activity forms a base-line of the students’ abilities to use CER before significant instruction in the framework.

It was my philosophy that students would best learn to use CER and internalize it for use on tests and other assignments if they learned it through practice and reflection. As a result, I integrated CER into every lab the students did throughout the study. Where I previously had students answer post-lab questions, I instead used the CER framework and lead a discussion about each one as a whole class. This also shortened the amount that students needed to write in the post-lab, focused it, and brought a consistency to our activities. After each activity, we discussed the best ways to support claims with evidence and worked on taking the extra step to provide reasoning.

In order to further develop the students' understanding of what makes an effective CER, I asked eighth-grade students to examine post-lab CERs that my current seventh-graders produced and later that their classmates had written. The eighth-grade students examined the CERs and worked together in small groups to decide what makes the best evidence and reasoning and to provide tips to the anonymous students as to how to improve their work. There were several reasons why I choose to have students critique other students' work instead of asking them to just improve their own. I found that students were much more open to give advice and make corrections on another student's work than they were to critiquing their own. It was easier for them to see the difference between a good reasoning and one that was lacking when pride was not in the picture. No data was collected from this process; it was used as a tool to improve the students' skills with CER.

During the eighth-grade treatment, all students took two tests, the Biomes and Ecosystems Unit test and a midterm, and as a result completed two After-Test Surveys (Appendix C). Not all of the students took the Biomes and Ecosystems test at the same time and only a handful of students had taken an additional assessment by the conclusion of the study so a third treatment portion assessment was not included. As a result, there is no post-treatment test data. For some students, there was two months between the Biomes and Ecosystems test and the end of the study. Data was collected from short answer questions on these assessments to determine if students were transferring their CER skills to testing situations. Students were not prompted to provide evidence and reasoning on tests. Occasionally the test questions asked students to explain their

answers, but I did not add any additional prompts to the original questions. Students were not graded on their use of evidence and reasoning on the tests unless the original question asked for an explanation.

During treatment, eighth-grade students completed five CER lab activities and three of these were coded for quality using the CER coding scheme (Table 1) (Appendix E). Some labs were not coded because they were done as a class or with significant teacher intervention so the coding did not accurately represent the students' independent abilities. Unlike the test score data, the CER labs were evenly spread throughout the three months of treatment.

Table 1
Eighth-Grade CER Lab Activities

Activity	CER Question(s)	CER Coding
1. Is air matter?	Is air matter?	No
2. Heartrate CER	How does exercise affect heart rate?	Yes
3. Viruses CER	Are viruses living things?	Yes
4. Plant Simulation	Three questions on the effect of various variables (soil, sunlight, water, fertilizer, etc.) on plant growth or germination. Different students experimented with different questions.	Yes
5. CER Practice	Which are better, cats or dogs? Who should be our next president? Should kids have gym every day?	No

To assess and track students' ability to use the CER framework, I coded their responses on labs and activities that used CER as well as short answer test questions from all of the eighth-grade tests. This CER Coding scheme allowed me to collect data on students' ability to use the CER framework when they were prompted and supported by their classmates and me, as well as the ability to transfer those skills to a testing situation

(Appendix E). Part of the coding process involved developing an “outcome space” to predict student responses in order to accurately and fairly code them (Brown, Nagashima, Fu, Timms, & Wilson, 2010) (Table 2). The coding served as a way to observe the change in students’ ability to use CER and the quality and frequency of their use of it in testing and lab situations.

Table 2

Outcome Space for How Does Mass Affect Friction?

Claim	<p>Fully Valid (FV) – Increasing the mass of an object increases the friction on an object</p> <p>Partially Valid (PV) – Increasing the mass affected the friction</p> <p>Invalid (IV) – Decreasing the mass of an object increases the friction on an object</p> <p>No Link (NL) – We added more mass; The friction changed; Mass affects friction</p> <p>No Claim (NC) – No Claim</p>
Evidence	<p>Sufficient Evidence (SE) – With 1kg the friction on the block was 3N, with 500g the friction was 2N and with 200g the friction was 1.5N; When there was the most mass the force was 3N and with the least mass it was 1.5N.</p> <p>General Evidence (GE) – When there was more mass on the block, there was a greater force of friction</p> <p>Insufficient Evidence (IE) – The block with 1 kg had a friction of 3N</p> <p>Unrelated Evidence (UE) – The block was on the table; the sandpaper caused the friction; I did the lab.</p> <p>No Evidence (NE) – No Evidence</p>
Reasoning	<p>Connects to scientific principles and connects evidence to claim (CC) – I know my claim is true because the greater mass resulted in a greater friction because increasing mass increases the force of gravity on the block and the surfaces interacted more.</p> <p>Connects to principles (CP) – I know my claim is true because if there is more mass gravity will pull the objects down more and cause more friction.</p> <p>Connects evidence to claim (CE) – I know my claim is true because the greater the mass the greater the friction force was.</p> <p>No connections made (NC) – The force meter showed me the force; I did the lab; my teacher told me.</p> <p>No reasoning (NR) – No reasoning</p>

Note: Blue color indicates mastery level codes.

There were multiple codes for evidence and reasoning that I considered mastery of the skill. For evidence, responses that were coded as “sufficient evidence” as well as “general evidence” showed that a student was able to provide effective evidence. A coding of “general evidence” is a slightly lower category than “sufficient evidence” because it just described a relationship or summarized the data instead of giving specific data points. However, providing “general evidence” is enough to adequately support claim. For a student’s reasoning, mastery level of the skill could be coded as “connects evidence to claim” or “connects to principles,” and when a student used both of these in their reasoning it was coded as “connects principles and evidence to claim.” Although “connecting to principles” shows a higher skill level than “connects evidence to claim,” both demonstrate proficiency in writing reasoning.

Three students in the eighth-grade class were also interviewed after treatment using the Post Interview Questions (Appendix F). The interview answers were analyzed for themes in student responses.

Seventh-grade Methodology

The seventh-grade classes began the year using CER. They participated in a variety of CER learning experiences throughout the year. Our work with CER was accompanied by Friday Skills activities focusing on graphing, independent and dependent variables, and relationships between variables. Towards the end of treatment, seventh-grade students participated in two CER practice activities that included identifying the parts of a CER and analyzing their classmates’ anonymous work. During the first activity, CER Practice about Cockroaches, students highlighted the claim, evidence, and

reasoning in sample responses from fictional students (Figure 3). Afterwards, the class discussed the best parts of each student's response and how they could improve. The second practice activity was almost identical to the ones that I used with my eighth-grade students and allowed the students to think about what makes an effective CER in a more comfortable environment than having to confront their own work.

CER Practice

Do cockroaches prefer light or dark environments?

Student 1:

Cockroaches prefer dark environments. They want to hide from us. Because cockroaches have those long antennae, maybe they can just feel their way around and don't need light to see.

Student 2:

Cockroaches prefer dark environments. When we let 10 cockroaches choose between the light and dark side of a box, 9 of them chose the dark side. Since more of the roaches chose the dark, we can tell that is what they prefer.

~ Claim
~ Evidence
~ Reasoning

Figure 3. CER Practice about cockroaches.

Throughout the school year, the seventh-grade students participated in ten CER lab activities (Table 3). Eight of their responses were coded with the CER Coding scheme (Appendix E).

Table 3
Seventh-Grade CER Lab Activities

Activity	CER Question(s)	CER Coding
1. Drops on a Penny	How does the amount of soap mixed with water affect how well the water stays on the penny?	Yes
2. Friction Lab	How do various variables (mass, surface area, or texture) affect the force of friction? Students chose which variable to investigate.	Yes
3. Ramp Lab	How does the steepness of the ramp affect the force needed to pull the block up the ramp?	No
4. Pendulum Lab	How do various variables (mass, length, or angle) of a pendulum affect its period? Students chose which variable to investigate.	Yes
5. Thermal Expansion Lab	What happens to the size of an object as it heats up?	Yes
6. Frozen Pucks Lab	How does the temperature of a hockey puck affect the height it will bounce?	No
7. Magnet Strength Lab	Is one pole of a magnet stronger than another?	Yes
8. Speed of Sound Lab	Does sound travel faster in solids or gases?	Yes
9. Is air matter?	Is air matter?	Yes
10. Density Simulation	How can you predict whether an object will sink or float? How can you tell if a crown is made of solid gold?	Yes

During the capstone project, my seventh graders took five tests, and I had collected CER coding data from the first one to test the coding scheme and serendipitously had that data to include in my capstone project. In addition to these test scores and one instance of CER coding, I collected additional coding data from a density quiz during the last week of treatment. Seventh grade students did not complete the After-Test Confidence Survey, but I interviewed nine students with the Post Interview

Questions (Appendix F). I choose five students from the advanced class and four students from the regular class to interview.

The data collection tools used to answer each of the research questions are outlined in Table 4.

Table 4
Data Triangulation Matrix

Research Component	Data Source	Data Source
Will purposeful teaching of CER skills and repeated use of CER in inquiry experiments and assessments improve student test scores traditional tests?	Test scores from unit tests	
Will it improve student confidence on tests?	After-Test Confidence Survey	Student Interviews
Will it improve student success with the CER technique?	CER Coding Scheme on labs and tests	Student Interviews

DATA AND ANALYSIS

Eighth Grade Data and Analysis

Results showed a slight improvement in eighth-grade test scores during the treatment period ($N=15$). The class average for the pre-treatment test, Ecology and the Environment, was 77% and the averages for the post-treatment tests, Biomes and Ecosystems and the Midterm, were both 79%. The comparative boxplots show a visible increase in the lower quartile and median (Figure 4). The student with the outlying score on the Ecology and the Environment test scored similarly on the Biomes and Ecosystems test until I decided to administer some testing accommodations, namely having the test

read to the student. Their score increased to a 77% but was not allowed accommodations on the Midterm and the student's score was a 49%.

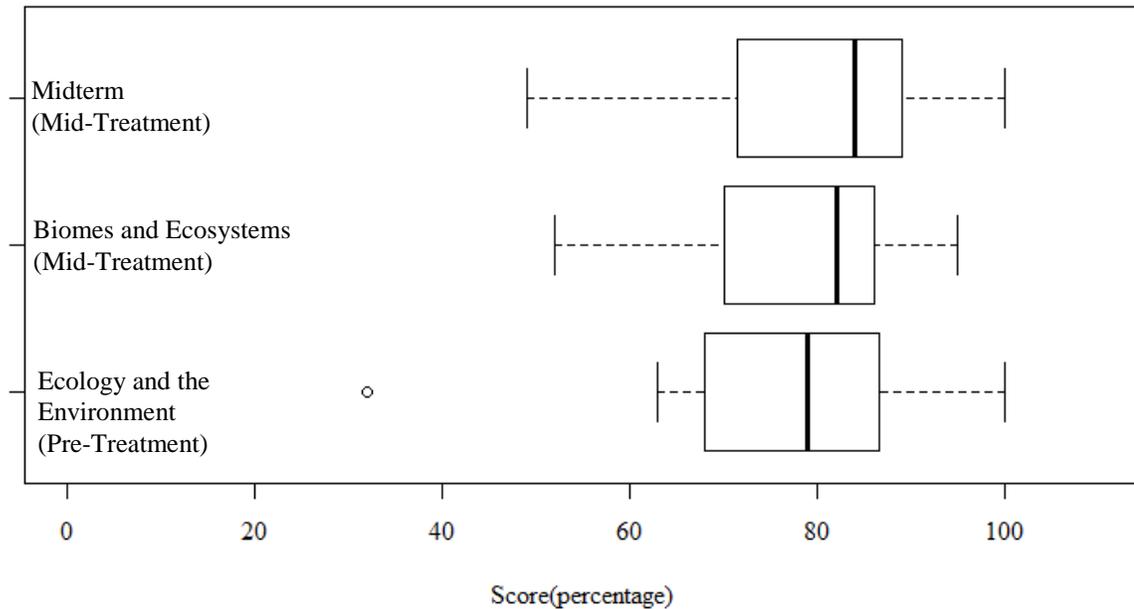


Figure 4. Comparative boxplots of eighth-grade test scores, ($N=15$).

Eighth-grade students reported their confidence on three aspects of test-taking (Appendix C). Continued use of CER did not improve students' self-reported confidence in their ability to understand test questions or be able to explain themselves on short-answer questions. The only notable improvement in confidence was in understanding between the Ecology and the Environment Unit and the Biomes and Ecosystems Unit (Figure 5). This coincides with the change in instruction through personalized learning.

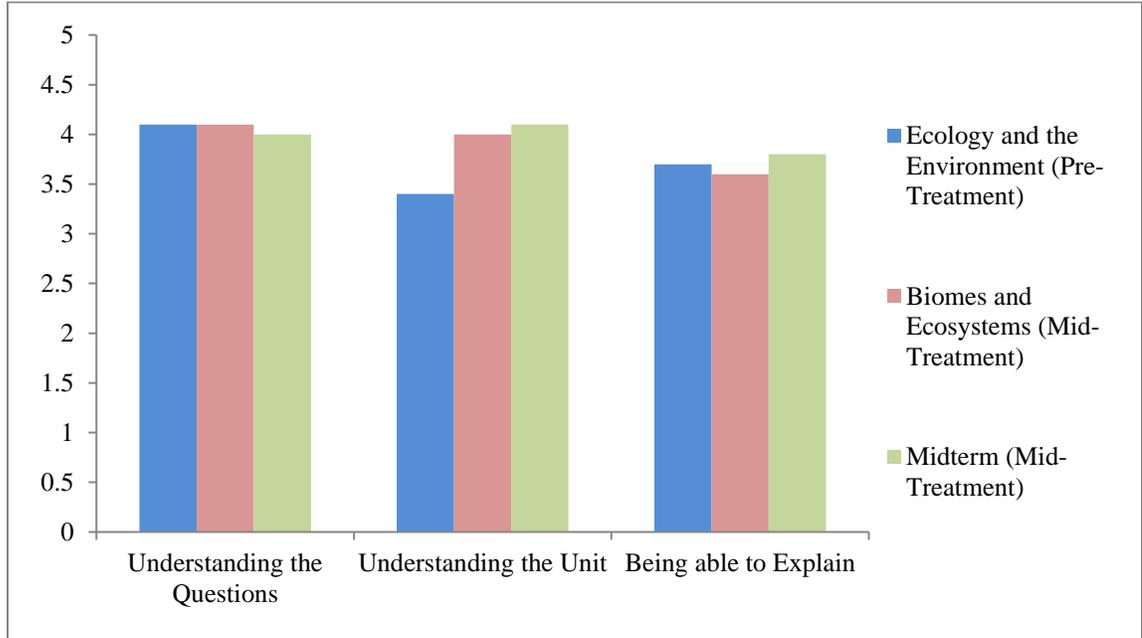


Figure 5. Average student confidence from the After-Test Survey, ($N=15$).
Note. 5=Strongly Agree, 4=Agree, 3=Neutral, 2=Disagree, 1=Strongly Disagree.

At the end of treatment, the twelve students who were interviewed gave added insights into student confidence on tests. When asked if multiple choice or short answer questions were easier on a test, ten students responded that they think multiple choice questions are easier. They had three reasons for feeling this way. Five students said multiple choice questions were easiest because you can rule out some answers, four said that you don't have to think about what to write, and one student said that it is easier because the possibilities aren't endless. The two students that preferred short-answer questions said that multiple choice questions are often trying to trick you.

Students were also asked about their confidence in their ability to explain themselves on a short-answer question and echoed the quantitative data. Six students reported feeling pretty confident, four reported feeling a medium or okay level of confidence, and two students were not very confident. One of the students who felt

pretty confident remarked that she is confident “because I can go over it if I’m not confident.”

Student success with the CER technique was tracked on both tests and labs. Eighth-grade students’ ability to both provide sufficient evidence and support claims with reasoning on labs improved during the treatment phase of the project (Figure 6). All of the students were able to write fully valid claims during the lab at the beginning of treatment as well as at the end of treatment. Ten percent of students wrote an invalid claim on the most difficult lab, the Virus CER.

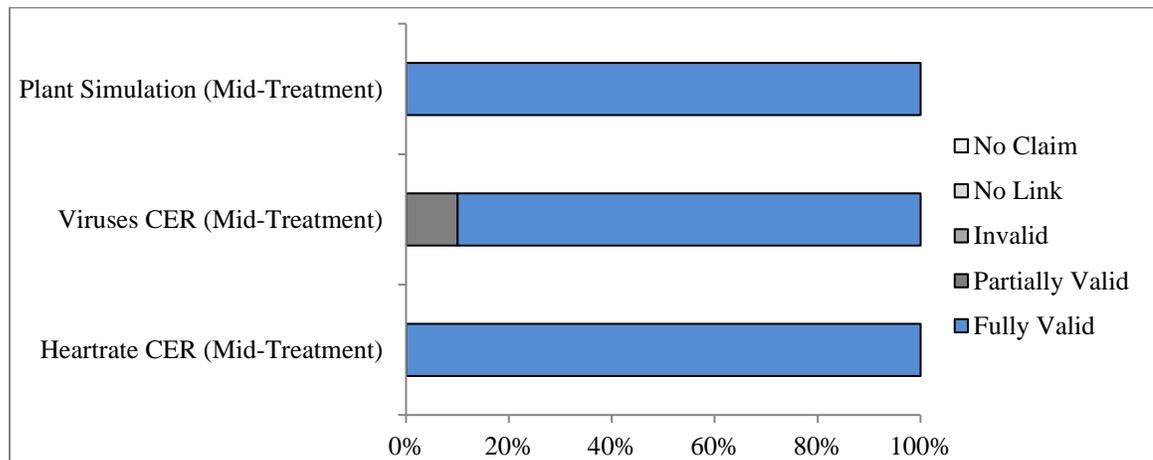


Figure 6. Eighth-grade coding for claims made on labs. All of these labs were done during treatment with Heartrate CER being the first and the Plant Simulation being the last, ($N=15$).

Three eighth-graders reported feeling, “pretty confident,” “very confident” and “pretty okay” about their ability to write a good CER after a lab. Two of them felt that the reasoning was the hardest part of this process because “sometimes I can’t come up with reasoning” and “if there isn’t a lot of evidence it can be hard.” The third student felt that the evidence was the hardest because you have to do the lab to get the data. When asked how they would address someone who was wrong about something they knew about

science they all said that they would look it up or tell the other person to look it up so that the other person could see the evidence.

Students greatly improved their ability to provide sufficient evidence from the beginning of treatment to the end of treatment. Sixty percent of students provided sufficient evidence on the first lab activity, Heartrate CER. Sufficient evidence increased to 70% on the second activity, Viruses CER, and finally to 78% on the Plant Simulation (Figure 7). An additional 9% of students provided general evidence on the plant simulation. Students were able to articulate that a good piece of evidence requires multiple specific details. One student said that it needs to have “lots of information and the data that supports the claim.”

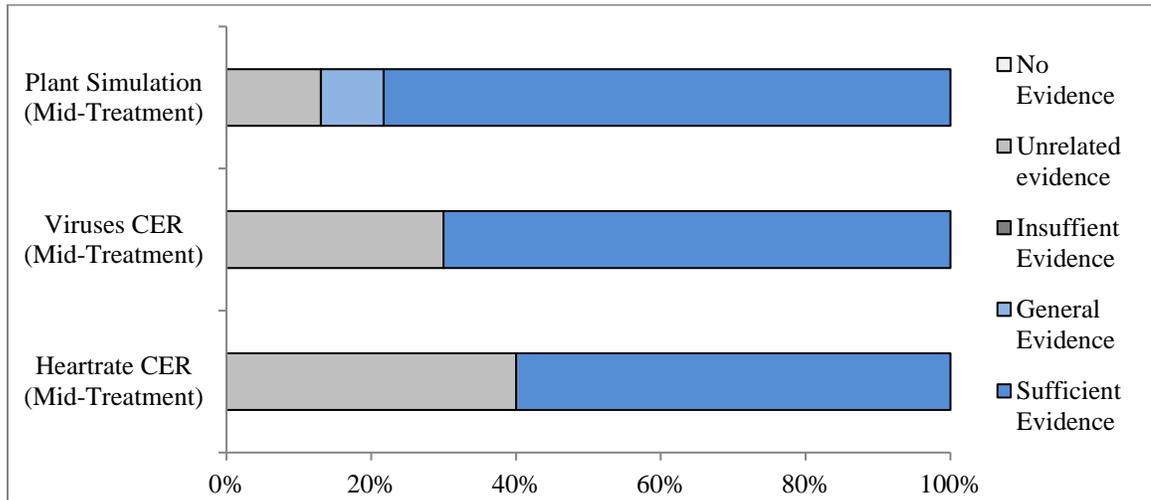


Figure 7. Eighth-grade coding for evidence on labs. All of these labs were done during treatment with Heartrate CER being the first and the Plant Simulation being the last, ($N=15$).

Note. Blue color is used to represent mastery and gray color represents a lower skill level.

Students also improved their ability to provide reasoning on lab activities. 10% more students made connections on the Virus CER than the Heartrate CER (Figure 8).

On the lab that followed, the Plant Simulation, 13% of students were able to use the

highest level of the reasoning skill by connecting to principles and connecting the evidence to the claim. An additional 30% of students were able to make other connections for a total of 43%. Students who were interviewed were asked what makes an effective reasoning. Their ideas included: making sense, using scientific language, using detail, and one student said that a good reasoning will “connect the evidence and the claim together.”

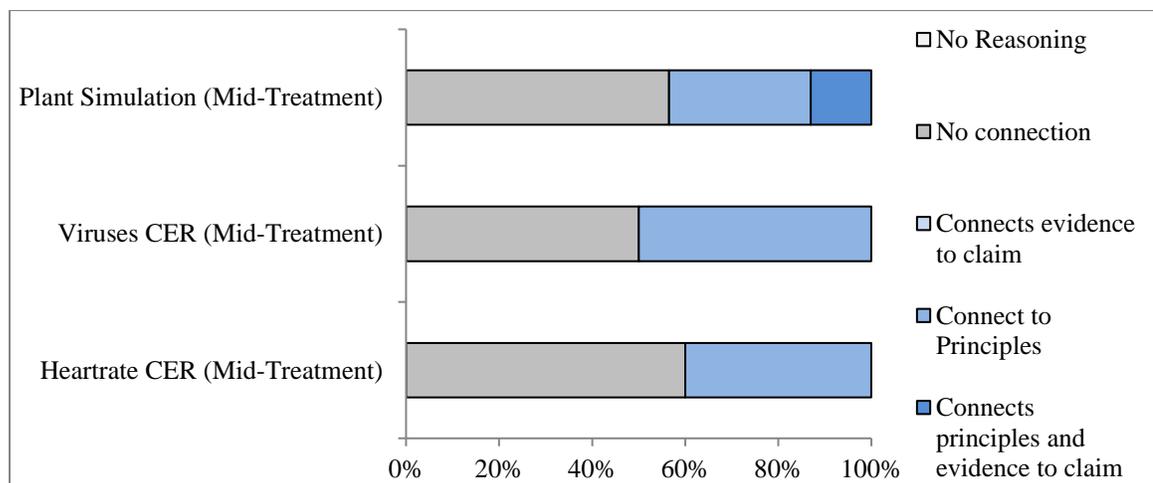


Figure 8. Eighth-grade coding for reasoning on labs. All of these labs were done during treatment with Heartrate CER being the first and the Plant Simulation being the last, ($N=15$).

Note. Blue color is used to represent mastery and gray color represents a lower skill level.

Students remained consistent in their ability to provide claims with 57% providing a fully valid claim on the Ecology and the Environment test and 58% on the Biomes and Ecosystems test (Figure 9). However, students who provided no claim did not necessarily get the question wrong, they may have made a statement of evidence or reasoning that answered the question. This coding data just tracks the frequency of students’ usage of the different parts of the CER framework.

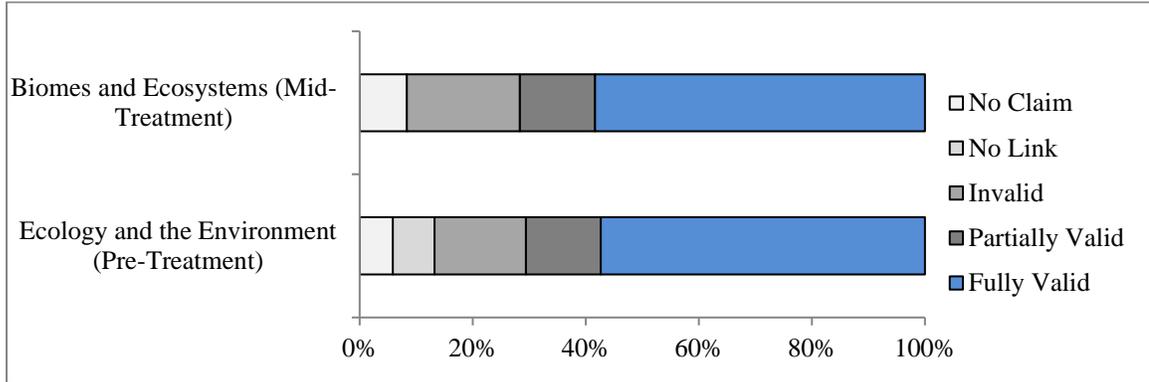


Figure 9. Eighth-grade coding for claims on tests, (N=15).

Students showed a decrease in their tendency to provide evidence and reasoning on assessments. The percentage of students providing general or sufficient evidence went from 26% to 17% (Figure 10). Reasoning significantly decreased from 25% of students initially providing a connection to principles to 4% of students providing any reasoning (Figure 11).

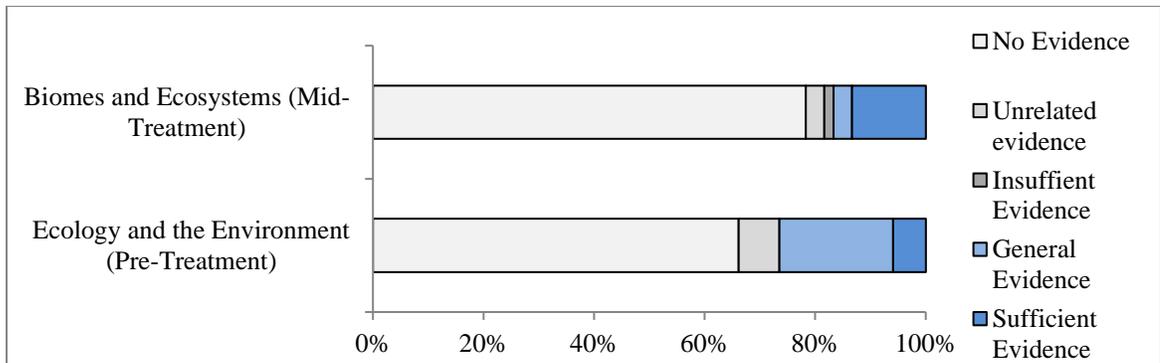


Figure 10. Eighth-grade coding for evidence on tests, (N=15).

Note. Blue color is used to represent mastery and gray color represents a lower skill level.

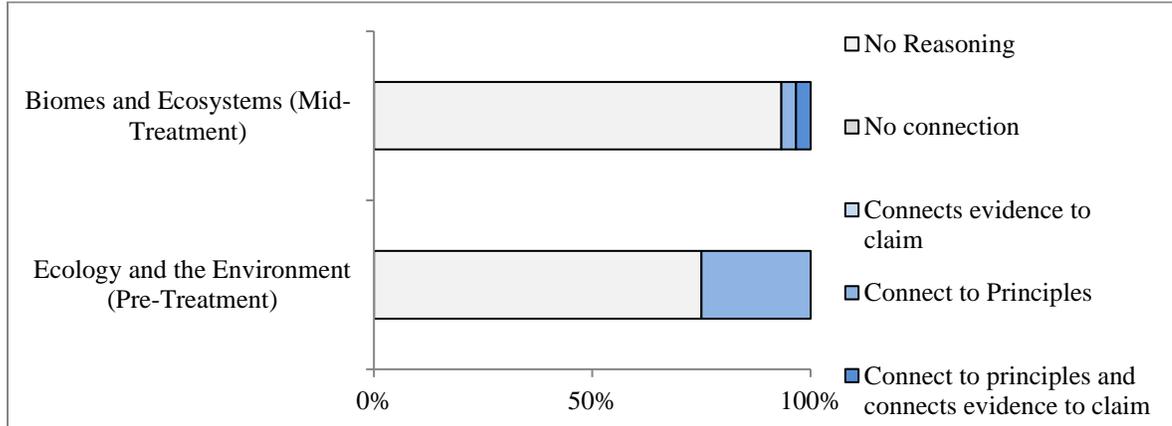


Figure 11. Eighth-grade coding for reasoning on tests, ($N=15$).

Note. Blue color is used to represent mastery and gray color represents a lower skill level.

Although no improvements were made in students' test scores, testing confidence, or argumentation skills on tests, eighth-grade students improved their CER skills on labs. By the end of treatment, the number of students who showed mastery level on providing evidence on labs increased by 27% for a total of 87% mastery (Figure 7). Students also improved their ability to write effective reasoning with the highest mastery percentage being 43% (Figure 8).

Seventh Grade Data and Analysis

Results showed that the use of CER throughout the school year did not improve seventh-grade test scores ($N=32$). The averages for the five tests were 76%, 80%, 78%, 82%, and 79%. Student scores were lowest on the first test, Forces and Motion, and comparative box plots show that test had a large lower quartile but no other difference are apparent (Figure 12).

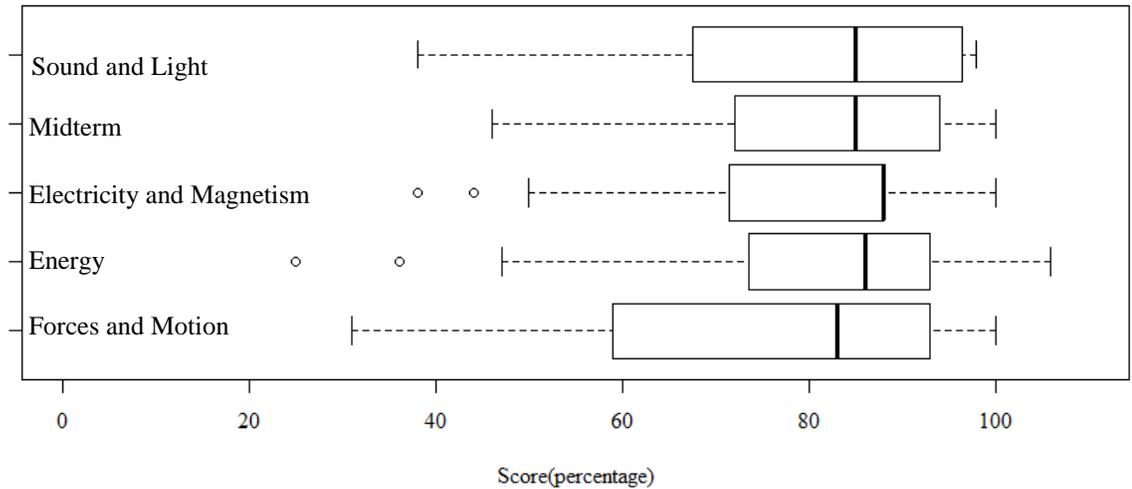


Figure 12. Comparative boxplots for seventh-grade test scores. All of these assessments were given during treatment with Forces and Motion being the first test and Sound and Light being the last, ($N=32$).

Seventh grade labs were also coded for CER proficiency. Students remained steady in their ability to provide evidence but their use of fully valid claims and the sophistication in their reasoning increased throughout treatment. Students' ability to write fully valid claims improved with the last five labs all having higher percentages of fully valid claims when compared to the first three labs (Figure 13). There is one lab, Thermal Expansion, where the fully valid claim must have been particularly clear to see or easy for students to determine because 100% of students provided a fully valid claim.

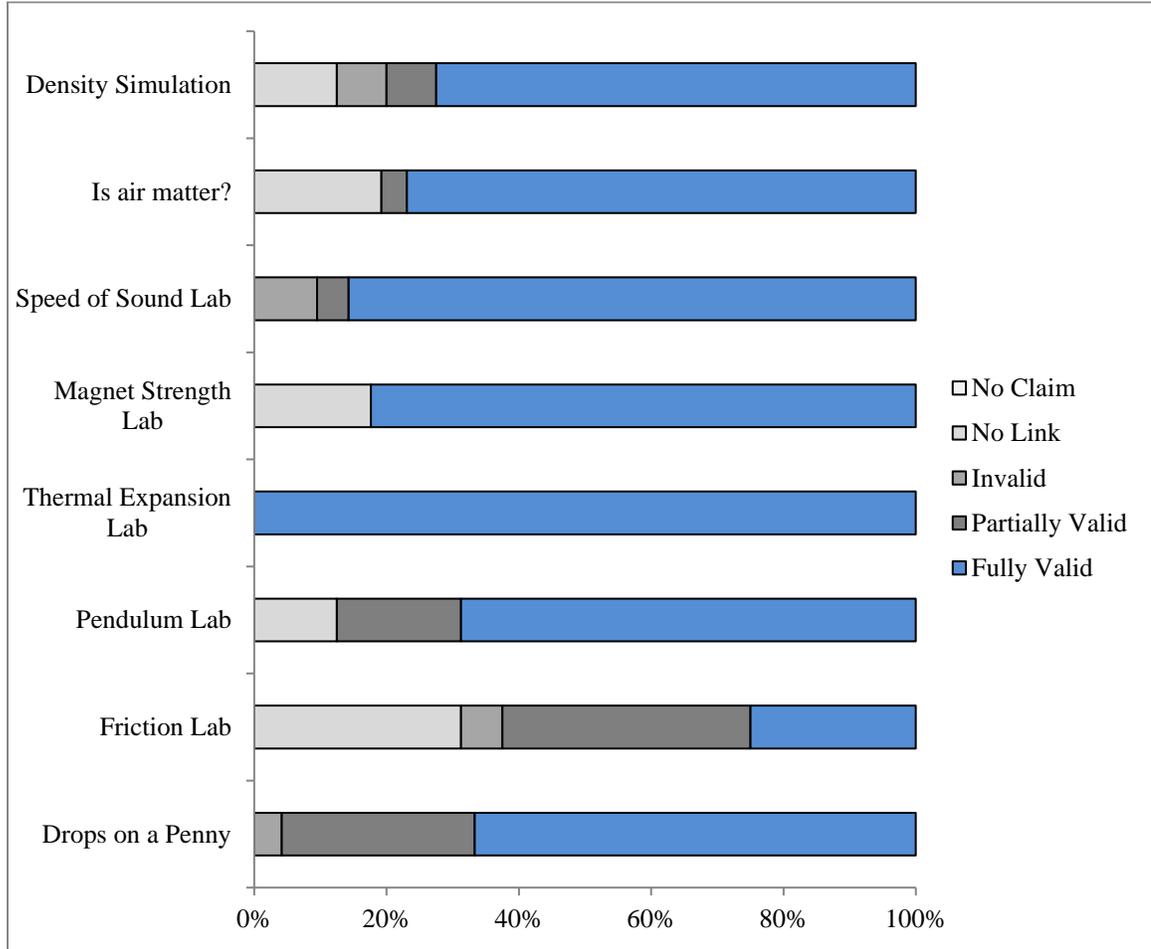


Figure 13. Seventh-grade coding for claims on labs. All of these labs were during treatment with the first being Drops on a Penny and the last being the Density Simulation, ($N=32$).

The nine interviewed students reflected on their ability to write a good CER. Five voiced that they felt very or pretty confident that they can write a good CER. Three had medium confidence and one was not confident at all. One student said, “I’m more confident if we discuss what we do with a group first” and another remarked that “it depends on the lab. I won’t be as confident if the lab’s confusing.”

When asked what part of a CER is the most difficult to write, three students thought it was the evidence, one student reasoning, and another student the claim. One

student said that nothing was difficult about it but that “it was hard in the beginning.”

Three more students, all from my advanced class, said that both evidence and reasoning were hard “because they’re pretty much the same thing.” These students had multiple methods for addressing a classmate who had a different idea about science than them.

Their ideas included: using reasoning, arguing until they make their point, using evidence, telling the other person to look it up, yelling, not putting the other person down, and waiting until the other person finds out for themselves.

Seventh-grade students did not improve their ability to provide evidence on labs but were consistent in their abilities. Students were able to provide sufficient or general evidence 60-70% of the time on six of the labs (Figure 14). The other two labs, Drops on a Penny and Thermal Expansion, had over 90% of students providing sufficient evidence. Six of the nine students interviewed said that it was important that evidence matches the lab that was done. Two students added that it is good to use numbers. Another two students thought the evidence should match the claim, and two more said that the evidence must be clear.

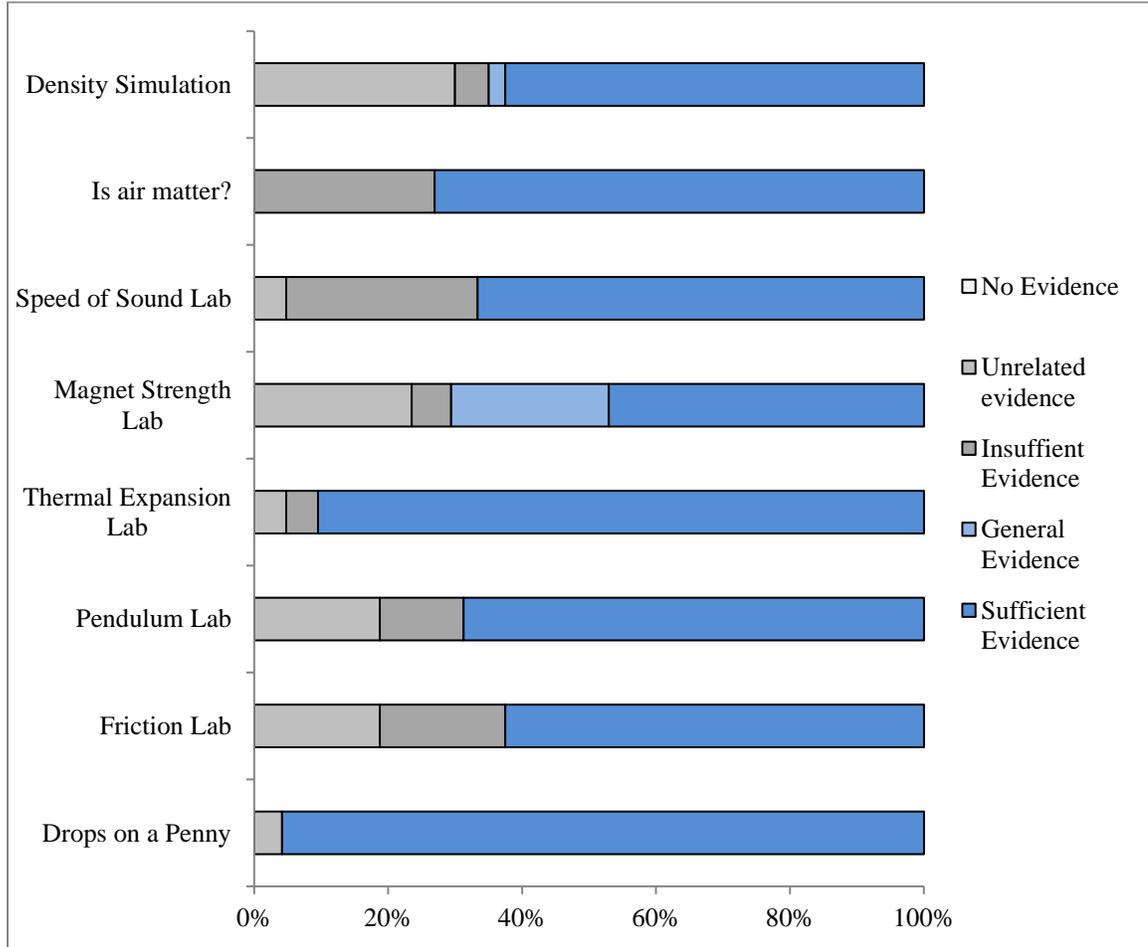


Figure 14. Seventh-grade coding for evidence on labs. All of these labs were during treatment with the first being Drops on a Penny and the last being the Density Simulation, ($N=32$).

Note. Blue color is used to represent mastery and gray color represents a lower skill level.

Seventh-grade students were consistent in their use of reasoning but their use of higher-level connections increased as treatment progressed. Eight of the labs had between 40 and 60% of students providing some sort of sufficient reasoning (Figure 15). The only lab with a percentage outside of that range was the second to last activity, Is air matter?, and 77% of students provided reasoning. Students were more likely to make sophisticated connections that included connecting to scientific principles as the treatment went on. The percentage of students making connections to principles

increased each lab except for the Magnetic Strength lab and the last lab, the Density Simulation. From the first lab to the second to last lab this percentage increased from 17% to 54%. When interviewed, four students said that a good reasoning includes scientific facts and two said that it needs to match the claim. Five students said that it needs to match the evidence with one student clarifying that “it’s not evidence.” One student didn’t know what made good reasoning.

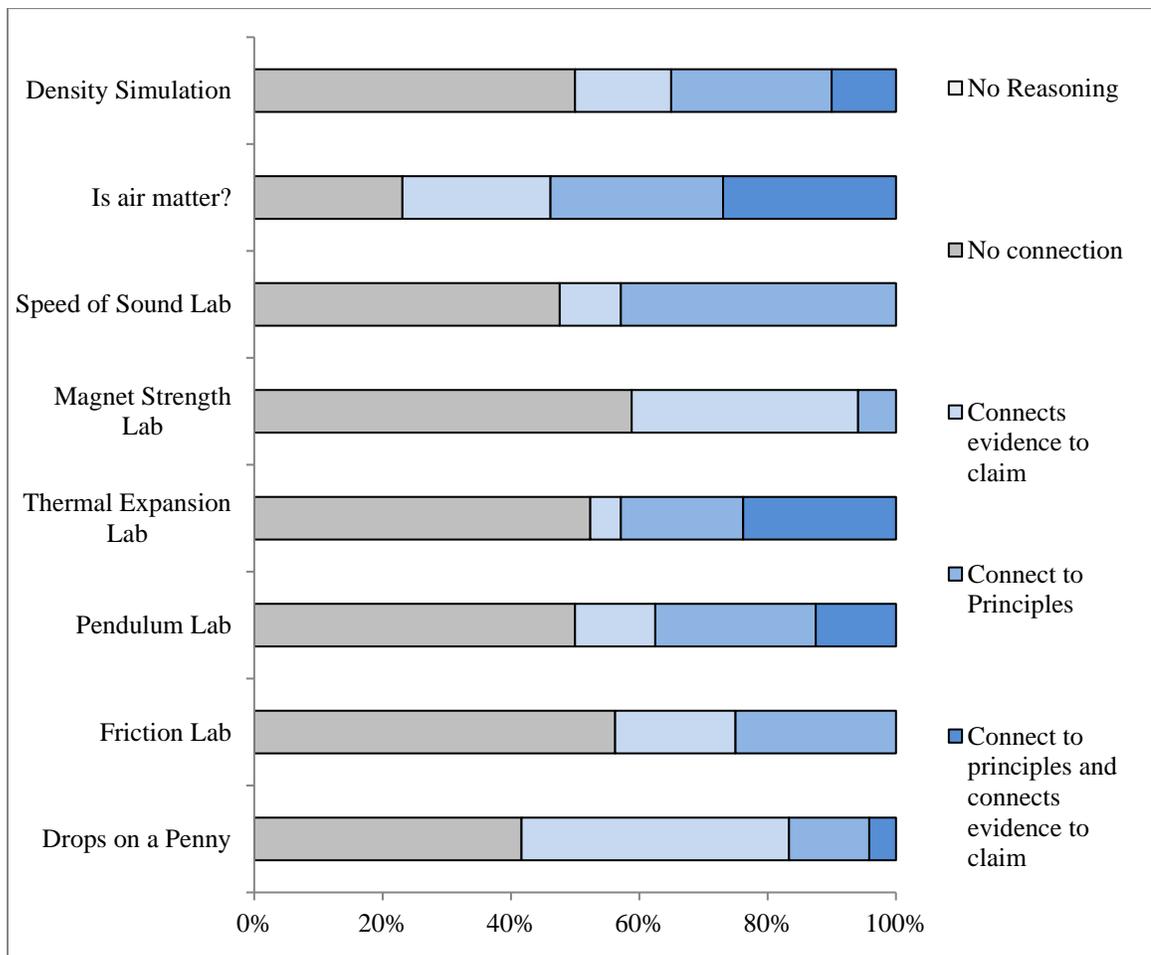


Figure 15. Seventh-grade coding for reasoning on labs. All of these labs were during treatment with the first being Drops on a Penny and the last being the Density Simulation, (N=32).

Note. Blue color is used to represent mastery and gray color represents a lower skill level.

The seventh-grade students drastically improved their tendency to use claims, evidence, and reasoning to answer test questions from mid-treatment to the end of treatment. Students increased their use of claims from 58 to 90% (Figure 16).

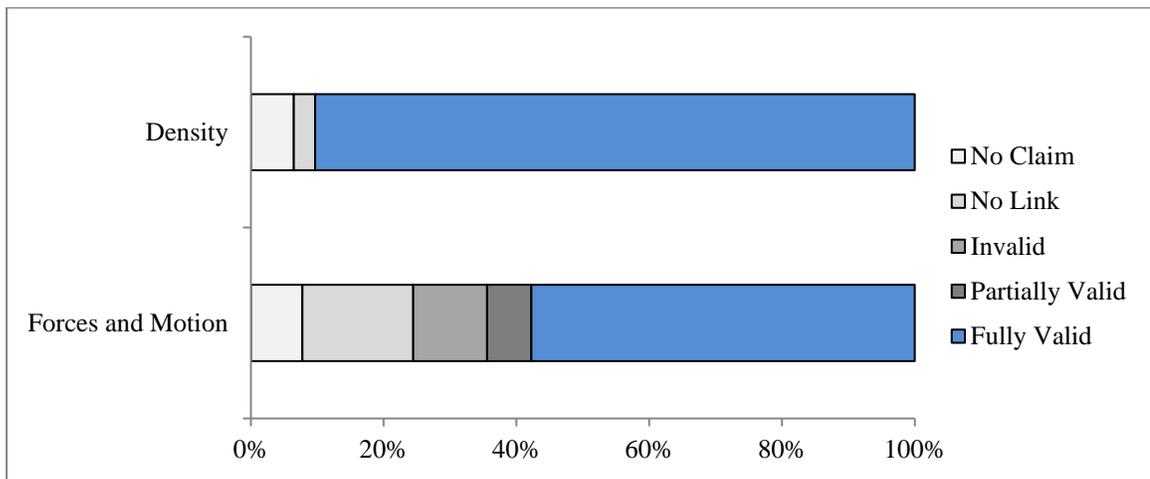


Figure 16. Seventh-grade coding for claims on assessments. The Forces and Motion test was given during treatment and the Density quiz was given post-treatment, ($N=32$).

Students increased their use of general or sufficient evidence from 17 to 80% (Figure 17). The one caveat in this is that the post-treatment assessment, the Density quiz, included questions where students did calculations and had data to include as evidence. However, this increase is substantial enough that I am confident that it shows a definite increase and not just fluctuations due to the type of question.

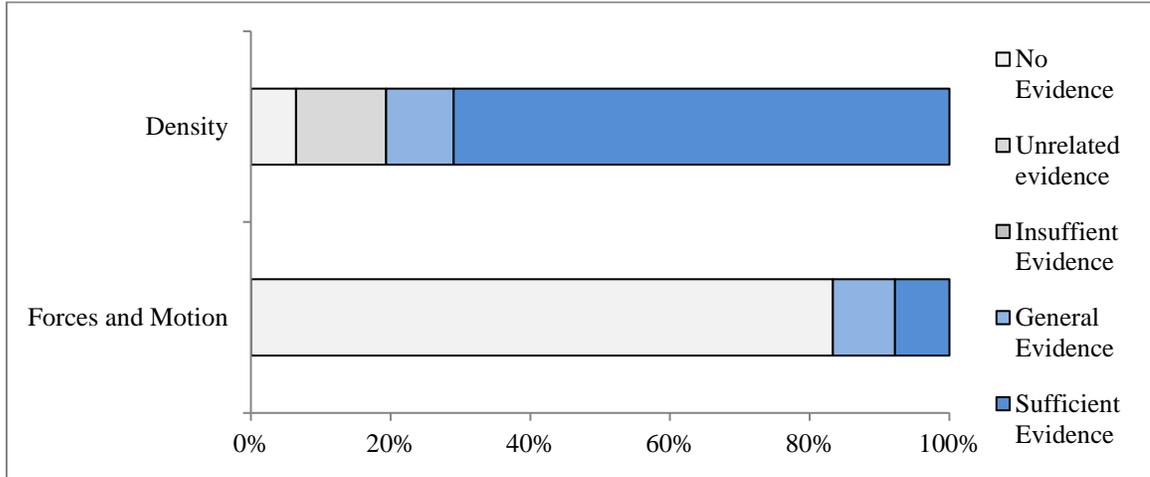


Figure 17. Seventh-grade coding for evidence on assessments. The Forces and Motion test was given during treatment and the Density quiz was given post-treatment, ($N=32$). *Note.* Blue color is used to represent mastery and gray color represents a lower skill level.

Students also increased their use of reasoning on assessments. More students provided reasoning on the post-treatment quiz that included a connection to scientific principles and connect the evidence to the claim than students who provided evidence that just included a connection to principles on the mid-treatment test (Figure 18). The percentage of students who included reasoning in their response increased from 40% to 84%. The percentage of students who wrote reasoning that included scientific principles increased from 40% to 65%.

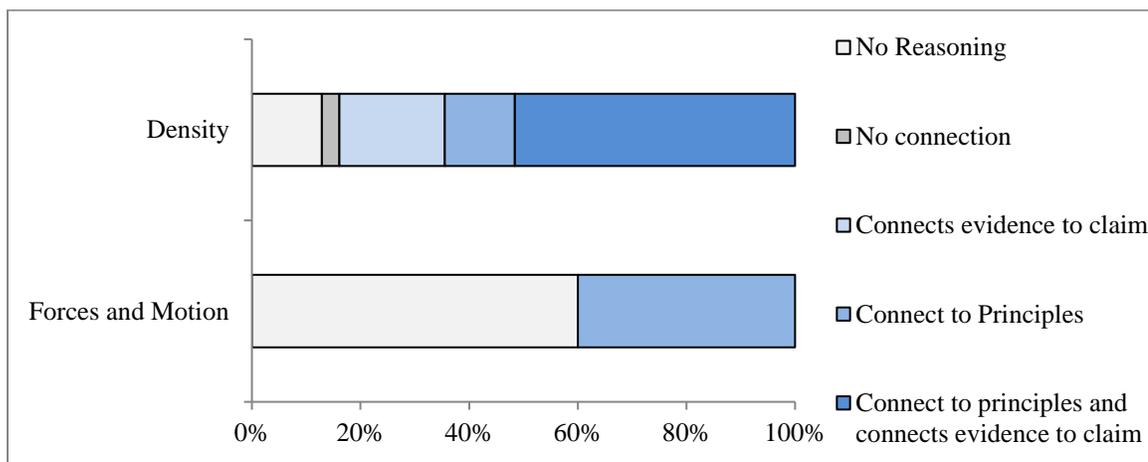


Figure 18. Seventh-grade coding for reasoning on assessments. The Forces and Motion test was given during treatment and the Density quiz was given post-treatment, ($N=32$). *Note.* Blue color is used to represent mastery and gray color represents a lower skill level.

Although seventh-grade students did not improve their test scores, they refined their use of the CER skill throughout treatment. By the end of this research project, 100% of students showed they could write a fully valid claim and 96% wrote CERs with sufficient evidence. Seventy-seven percent of students reached mastery level of writing reasoning and the percentage of students connecting their claims to scientific principles increased from 17% to 54%. Students also made progress in their use of argumentation skills on tests with 90% writing fully valid claims, 80% providing a mastery level of evidence, and 65% connecting scientific principles to their claims.

This improvement in CER skills is further supported by student work samples. The samples selected from one student's work from the beginning, middle, and end of data collection are representative of the progress made by many students in my classes. On the Friction Lab CER, this student, Student A, struggled with the CER process (Figure 19). Their claim did not answer the question and was coded NL for no link to the question. Their evidence was incomplete because they only used a small spring scale that

could not measure all of their data and was coded IE for insufficient evidence. Their reasoning is based on the incomplete evidence and was coded NC for no connections made. The sample taken from Student A's work about half way through treatment shows frustration with the CER process (Figure 20). This CER was coded NL (no link to question), NE (no evidence), and NC (no connections made). The third piece of student work, also from Student A, shows improvement in all three areas of the CER technique (Figure 21). Their claim was coded FV (fully valid), even though they just said "yes," this still correctly answers the question. This student did not do a good job of explaining their experiment and was given a code of IE (insufficient evidence) for this reason. The experiment description was important because matter has both volume and mass. The experiment description should have mentioned that the balloon was blown up, thus providing evidence that air has volume. However, Student A's measurements were accurate and included units which were an improvement over the Friction Lab. This student's reasoning in the last sample is a great example of connecting to principles (CP). They took the information they knew about matter and connected that to their claim that air was matter. The only way this student could have improved their reasoning would be to have added a comment about how their evidence showed the air had both mass and volume.

Claim: Your claim is the answer to your question.
The mass station went all almost over

Evidence: Your evidence is the data that supports your claim.
100 is 1.5
300, went over
500, went over
1000, went over

Reasoning: Your reasoning describes how your evidence supports your claim. It often includes the scientific principles that made your claim and evidence true.
we got the mass all the way over the thing on almost everything

Figure 19. Student A's CER response on the Friction Lab. This student was responding to the question "How does mass affect friction?"

Claim: Your claim is the answer to your question.
The magnet worked

Evidence: Your evidence is the data that supports your claim.
It worked

Reasoning: Your reasoning describes how your evidence supports your claim. It often includes the scientific principles that made your claim and evidence true.
It worked like a magnet

Figure 20. Student A's CER response on the Magnetism Lab. This student was responding to the question "Is one pole of the magnet stronger than the other?"

Claim:	
Evidence:	
Explain the experiment -	
Mesuring mass of ballon	
Mass before air was added	<u>2.8g</u>
Mass after air was added	<u>3.1g</u>
Mass of air	<u>0.3g</u>
Reasoning:	Yes
The ballon does have matter because it has mass and volume.	

Figure 21. Student A's CER on the lab called "Is Air Matter? This student was responding to the question.

INTERPRETATION AND CONCLUSION

This study was designed to investigate if the intentional and continued use of the claim, evidence, reasoning (CER) technique would improve students' test scores, testing confidence, and overall ability to use the skill. This project was unable to show an improvement in test scores or testing confidence as a result of the use of CER in the science classroom. The use of CER, however, improved the quality of short answer responses that seventh-graders made on assessments, but the same result was not shown in eight-graders. Lastly, students developed a greater ability to use the CER technique when prompted to do so in lab situations.

Over the course of the project, students' test scores were remarkably consistent. I chose this project because I was looking for another skill that would help my students better represent what they know on a test. Although, the logical process of CER was beneficial to my students' learning and cognitive processes, it did not make the impact on testing that I hoped. This indicates that the standardized tests that I took test questions from were not effectively testing for analytical thinking skills and did not allow my students to use the CER tools to improve their test scores. Test averages around 80% are excellent, but I know that average means that there are a couple of students who scored close to 100% and several students who scored 60% or lower. I still sit with these low scoring students to go over their tests and can see that know more than 60% of the material and I will continue to look for new ways to help them.

I also did not see any notable improvements to testing confidence. Any confidence increase was probably due to personalized learning and having another assessment at the end of the eighth-grade treatment would probably emphasize that more.

The use of CER improved the quality of students' short answer responses on tests. This is very clear for the seventh-grade students and less clear for the eighth-grade students. However, I am confident that it improves their skill level because the seventh-grade assessment that showed this improvement was post-treatment and the eighth-grade assessment was mid-treatment. This development is more valid when seen through the lens of the labs that students performed. Seventh-grade students participated in lab activities that directly related to questions they would see on tests. For example, we performed a lab on thermal expansion and there was a test question on thermal

expansion. They were able to use the claim, evidence, and reasoning they investigated in the lab to prove their knowledge on the assessment. The only seventh-grade labs for which this didn't at least partially apply are the Drops on a Penny and Magnet Strength labs. The use of CER in labs directly supported the way students internalized the evidence for and scientific principles behind what they were learning and helped them think about it in a logical way. Students were able to transfer this to a testing situation.

Students' ability to use the CER technique in labs was greatly improved. I appreciated the consistency it gave to our labs. The students learned to expect that after they conducted their experiment they would need to review the question the lab was asking and figure out how their experiment answered that question. Over time, I think it began to help them think in a logical way about their lab results and this showed in the improvement in their CER writing. Interviewed students were overall confident in their ability to write a CER at the end of a lab and liked the support of a group. Based on their descriptions, they also had a good grasp on how to execute the three parts of the skill. However, it is clear that there is still work to be done and it was surprising that some of my advanced students felt that evidence and reasoning were the same. The interviews in which these comments were made were done during the last week of this capstone project. It would have been beneficial to have interviewed the students earlier so I could have caught and corrected this misconception before it was more deeply rooted. Regardless, we spent some time in the following weeks clarifying what a reasoning was and hearing from my students helped me focus my language and improve the way I caught CER. As these comments suggest and the results from this study confirm that

reasoning is the most difficult step of the CER process. I expect my students to continue to improve in their ability to write this part with continued practice, coaching, and support from groups.

Another interesting aspect of the results of this study was that the seventh graders had a bigger improvement and higher final percentages than eighth graders. There are several explanations for what may have caused this difference. The seventh-grade group has a larger number of high-level students because the “advanced” eighth graders are accelerated into ninth-grade science and were not a part of my study. However, studies have shown that the use of the CER technique is often most helpful to students with average, not high, test scores (Wang & Buck, 2015). The largest contributing factor was likely the seventh-grade group started with CER from the beginning of the year, instead of waiting to have a non-treatment phase. They were able to participate in many more labs and received feedback from me and their groups over three quarters of the school year instead of just one quarter. This shows that practice through labs, feedback, and group work over a long period of time was a more effective method of introducing CER than having students analyze their work and other students’ work like I did in the short time I had to work on CER with the eighth-grade group.

VALUE

Throughout this process and especially with the number of labs I was able to incorporate the claim, evidence, reasoning (CER) technique into with seventh-graders; I realized how much I love using it with labs. I saw a huge difference in students’ ability to synthesize the meaning of lab results compared to my previous method of post-lab

questions. It streamlined the way my students learned from labs and helped things make sense. It was also much easier to see if students understood the results or not because it is difficult to fake a good reasoning. This is contrasted with post-lab questions which were often easy to write shoddy answers for and get by. I will continue to incorporate the technique into as many labs as possible for my middle school students.

I also loved the discussions that it brought up in lab groups. The students that I saw the most improvement in and that got the most out of labs were the ones that spent time with their groups deciding on the best bits of evidence and coming up with scientific explanations for their claims. This is an aspect of the CER process that I plan on intentionally fostering in future years and to continue to build in my current students, especially the ones that I will continue to have next year.

I was thrilled when I read the responses my students were writing on their last few seventh-grade tests. Their explanations were not always perfect and not every student included a claim, evidence, and reasoning for each question, but they began to understand that they should and were able to show what they knew by adding evidence and reasoning. I am excited to see what will happen after using it so consistently with my seventh-grade students as they progress into their eighth-grade year and we will have more time to work on the skill together.

My plans to continue my work with CER include continued integration in the lab setting. I have also found that students like being surveyed because it gives them an opportunity to give me feedback on the changes I am making. I plan to find more labs that I can fit into my eighth-grade curriculum that the CER technique can be added to.

The next level of using CER in my classroom is to build a better culture for the use of argumentation. Students need to be able to not be afraid to voice their ideas and I need to give them opportunities to try more argumentation. If I interviewed my students at the end of next year, I want them to have had experience with effective argumentation skills so no one would consider “yelling” an appropriate response to someone who doesn’t agree with what they know about science.

REFERENCES CITED

- Aydeniz, M., & Bilican, K. (2016). Appropriating epistemic norms of science through sustained practice with argumentation: Can it happen? A learning progressions perspective. *Journal of Education in Science, Environment and Health*, 2(2), 125-133.
- Berland, L. K. (2011). Explaining variations in how classroom communities adapt the practice of scientific argumentation. *Journal of the Learning Sciences*, 20(4), 625-664.
- Brown, N. J., Nagashima, S. O., Fu, A., Timms, M., & Wilson, M. (2010). A framework for analyzing scientific reasoning in assessments. *Educational Assessment*, 15(3-4), 142-174.
- Bulgren, J. A., Ellis, J. D., & Marquis, J. G. (2014). The Use and Effectiveness of an Argumentation and Evaluation Intervention in Science Classes. *Journal of Education and Technology*, 23(1), 82-97.
- Chen, Y.-C., Hand, B., & Park, S. (2016). Examining elementary students' development of oral and written argumentation practices through argument-based inquiry. *Science and Education*, 25, 277-320.
- Johns, S., & Wolking, M. (2018). *The Core Four of Personalized Learning: The Elements You Need to Succeed*. Retrieved from Education Elements: https://www.edelements.com/hubfs/Core_Four/Education_Elements_Core_Four_White_Paper.pdf
- McNeill, K. L. (2011). Elementary students' view of explanation, argumentation, and evidence and their ability to construct arguments over the school year. *Journal of Research in Science Teaching*, 48(7), 793-823.
- Meacham, B. (2017, February 13). *Implementing the Claim, Evidence, Reasoning Framework in the Chemistry Classroom*. Retrieved from Chemical Education Xchange: <https://www.chemedx.org/article/implementing-claim-evidence-reasoning-framework-chemistry-classroom>
- NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: National Academies Press.
- U.S. News and World Report. (2014). *Romulus Central School*. Retrieved from Best High Schools: www.usnews.com/education/best-high-schools/new-york/districts/romulus-central-school-district/romulus-central-school-14027
- Wang, J., & Buck, G. (2015). The relationship between Chinese students' subject matter knowledge and argumentation pedagogy. *International Journal of Science Education*, 37(2), 340-366.

- Weir, I. (n.d.). *Pearson's correlation*. Retrieved from Statstutor:
<http://www.statstutor.ac.uk/resources/uploaded/pearsons.pdf>
- Weir, I. (n.d.). *Spearman's correlation*. Retrieved from Statstutor:
<http://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf>
- Windschitl, M. (2008). What is inquiry? A framework for thinking about authentic scientific practice in the classroom. In J. Luft, R. L. Bell, & J. Gess-Newsome (Eds.), *Science as inquiry in the secondary setting* (pp. 1-20). Arlington, Virginia: National Science Teachers Association.

APPENDICES

APPENDIX A
IRB EXEMPTION



INSTITUTIONAL REVIEW BOARD
For the Protection of Human Subjects
FWA 00000165

960 Technology Blvd. Room 127
 c/o Microbiology & Immunology
 Montana State University
 Bozeman, MT 59718
 Telephone: 406-994-6783
 FAX: 406-994-4303
 E-mail: cherylj@montana.edu

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

MEMORANDUM

TO: Miriam Harms and John Graves
FROM: Mark Quinn *Mark Quinn Ch*
 Chair, Institutional Review Board for the Protection of Human Subjects
DATE: November 27, 2017
RE: "The Effect of Claim Evidence Reasoning Instruction on 8th Grade Test Confidence and Test Scores"
 [MH112717-EX]

The above research, described in your submission of November 22, 2017, 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:

- (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.
- (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.
- (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.

APPENDIX B

ECOLOGY AND THE ENVIRONMENT TEST

Name _____

Science 8 Test

Ecology and the Environment

NOTE: These questions have been taken from previous NYS Science 8 State Tests

1. _____ A fox with thick fur would have a survival advantage over other foxes if
 - (1) there is more competition for food in the fox population
 - (2) the air temperature significantly decreases in winter
 - (3) a drought occurs, limiting the amount of water available
 - (4) a new disease appears that infects the foxes

2. _____ Competition within a pack of wolves may increase if there is an increase in the
 - (1) amount of food available
 - (2) amount of oxygen available
 - (3) size of their population
 - (4) size of the area they inhabit

3. _____ All of the different organisms interacting in a pond make up
 - (1) a community (3) water cycle
 - (2) a population (4) the habitat

4. _____ Which factor is most likely to cause the number of rabbits living in an area to increase?
 - (1) less water
 - (2) lack of shelter
 - (3) fewer predators
 - (4) limited food

5. _____ The information below describes a relationship between a type of fish known as a goby and a species of blind shrimp.
The shrimp digs a hole that provides shelter for itself and the goby fish. The goby fish lives at the opening to the hole and watches for predators. When a predator swims by, both organisms quickly move farther into the hole for safety. The relationship between these two organisms is best described as

(1) competitive	(3) beneficial
(2) harmful	(4) predatory

6. _____ Which type of relationship exists when a certain type of tree's roots need a fungus present in order to grow normally?

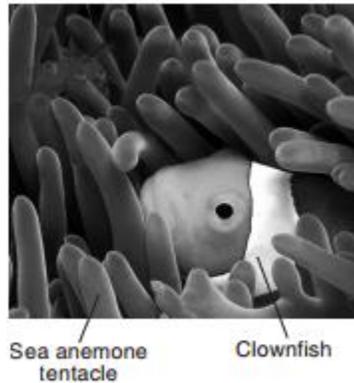
(1) beneficial	(3) harmful
(2) competitive	(4) infectious

7. _____ What eventually happens to a species when the environment changes and the organisms do not adapt?

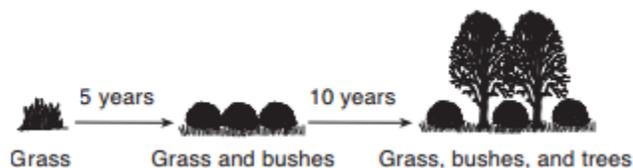
(1) classification	(3) extinction
(2) development	(4) regulation

Base your answers to questions 8 and 9 on the information below about two animals, the sea anemone and the clownfish, and on your knowledge of science.

Clownfish are tiny, omnivorous fish that find shelter from predators in the poisonous tentacles of sea anemones. The sea anemones sting their prey to capture food, but the clownfish are not hurt by the stinging tentacles. The clownfish clean the tentacles of the sea anemone and scare off butterfly fish, which consume sea anemones.



8. _____ The relationship between the sea anemone and clownfish is best described as
 (1) competitive (3) predatory
 (2) beneficial (4) harmful
9. _____ The clownfish is classified as an omnivore because it eats
 (1) both plants and animals
 (2) neither plants nor animals
 (3) only plants
 (4) only animals
10. _____ The diagram below represents changes in the main types of plant species found in a specific area over a 15-year period. Which process is represented in the diagram?



- (1) water cycle (3) environmental degradation
 (2) feedback system (4) ecological succession

11. _____ The data table below shows the yield of vegetables in a school's garden for 3 years. The yield is the number of pounds of vegetables harvested. The same number of plants was planted each year for all five vegetables.

Data Table

Vegetable	Yield per Year (pound)		
	2004	2005	2006
acorn squash	139	143	52
beet	93	122	81
butternut squash	147	103	30
onion	143	134	83
spinach	102	137	0

What is the most likely reason for the decrease in the vegetable yield in 2006?

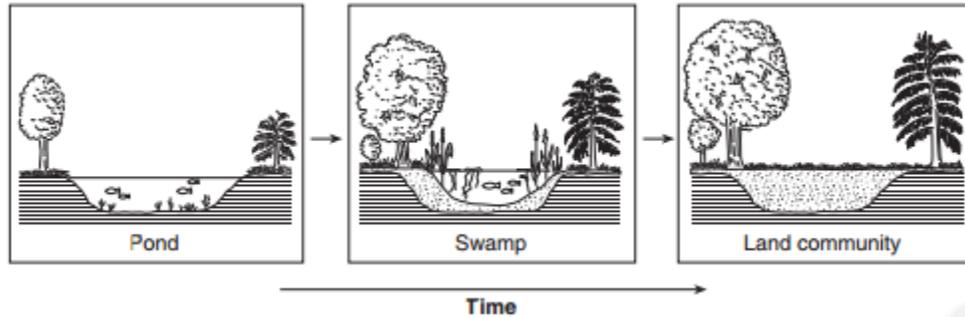
- (1) an increase in the size of the garden area
 - (2) an increase in the amount of sunlight
 - (3) a decrease in the rabbit population near the garden
 - (4) a decrease in the average yearly rainfall
12. _____ The drawings below represent three different birds. The beak shape makes it easy for each bird to obtain food in a different way. The differences in beak shape are examples of



(Not drawn to scale)

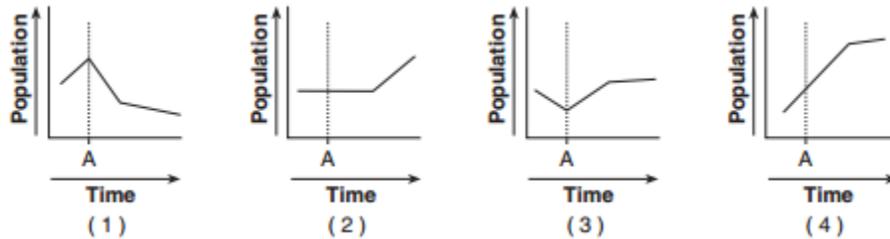
- (1) camouflage
- (2) competition
- (3) dynamic equilibrium
- (4) biological adaptation

13. _____ The diagrams below represent the same location over a period of many years.

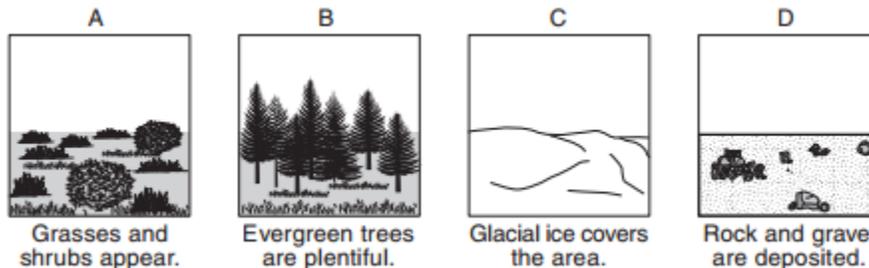


The sequence of diagrams best shows that, over time,

- (1) erosion increases
 (2) climates get colder
 (3) communities stay the same
 (4) ecological succession occurs
14. _____ Which graph shows what most likely would happen to the population of a certain animal if a new predator were introduced at time A?



15. _____ A glacier in Alaska has melted back a distance of 100 kilometers over the last 200 years. Four stages in this process are shown in diagrams A, B, C, and D below.

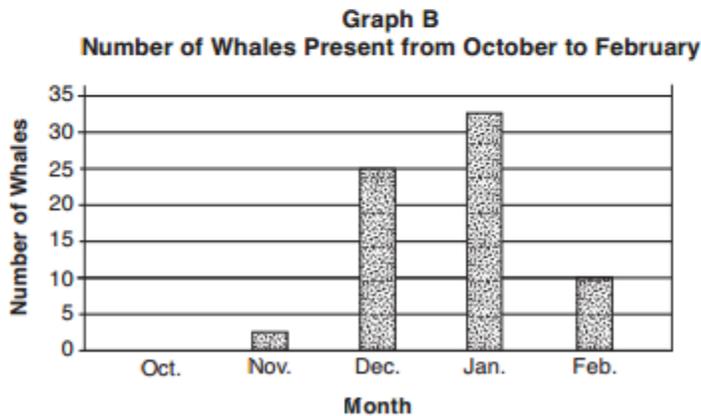
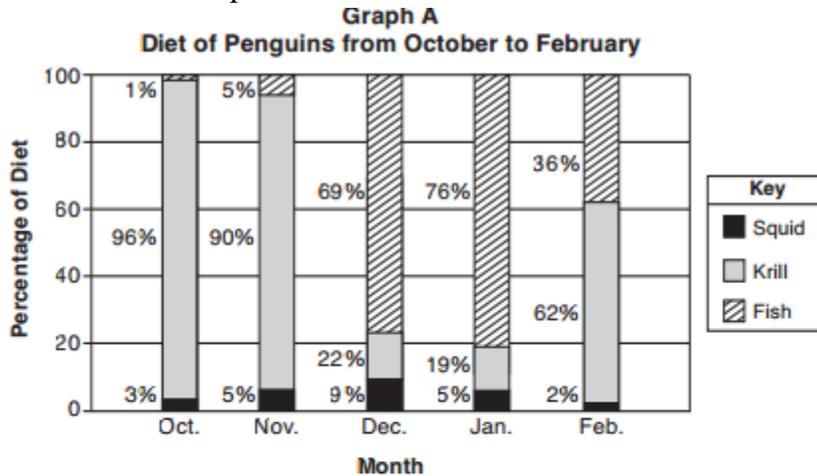


In which order should the diagrams be placed to represent the ecological succession that has taken place in the area?

- (1) C → D → A → B
 (2) C → D → B → A
 (3) D → C → B → A
 (4) D → C → A → B

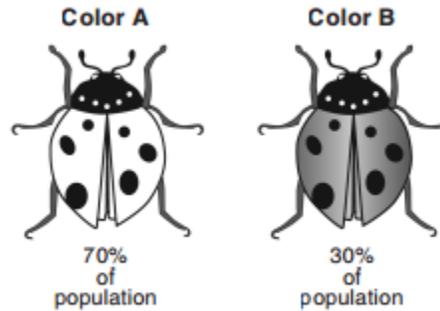
Base your answers to questions 16 through 18 on the bar graphs below and on your knowledge of science.

Graph A shows the diet of a particular species of Antarctic penguin over a five-month period. Its diet includes squid, krill, and fish, which are all animals found in the ocean. Graph B shows the number of whales that were spotted in the penguins' habitat during the same five-month period.



16. Describe one change in the diet of penguins when the number of whales present increased.
17. Based on the information on the graphs, give one possible explanation as to why the penguins' diet changed when the number of whales increased.
18. Explain why the penguins of this species are considered carnivores.

19. The diagram below represents a species of beetle (ladybug) with two different body colors labeled A and B.

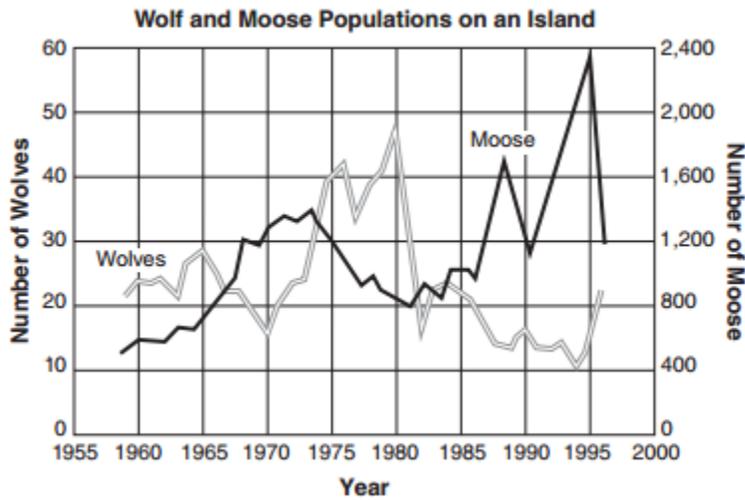


These beetles live on trees and are eaten by birds. The percentage of each body color in the population of this species is indicated. The habitat of this beetle population is a group of trees with light-colored bark.

Based on the information provided, explain why the beetle population in this habitat contains a higher percentage of beetles with body color A.

Base your answers to questions 20 and 21 on the graph below and on your knowledge of science.

The graph shows wolf (predator) and moose (prey) populations on an island over a period of many years.



20. Describe the changes in the wolf population from 1965 to 1970, and from 1970 to 1975.
- 1965 to 1970: _____
- 1970 to 1975: _____
21. Explain how the graph shows that there is a predator/prey relationship between the wolves and moose on this island.

APPENDIX C
AFTER-TEST SURVEY

AFTER-TEST SURVEY

For each of the following statements, select how much you agree or disagree with the statement.

1. I felt confident about my ability to understand the questions during this test.
Strongly Disagree Disagree Feel neutral Agree Strongly Agree

2. I felt confident about my understanding of the unit during this test.
Strongly Disagree Disagree Feel neutral Agree Strongly Agree

3. I felt confident about my ability to explain myself on this test.
Strongly Disagree Disagree Feel neutral Agree Strongly Agree

4. Complete the following sentences about this unit.
I feel confident because...

I wish I was more confident about...

5. Is there anything else you want me to know about the test?

APPENDIX D
IS AIR MATTER?

Is air matter?

Claim:

Evidence:

Explain the experiment –

Mass before air was added _____

Mass after air was added _____

Mass of air _____

Reasoning:



APPENDIX E
CER CODING

CER CODING

Based on techniques from “A framework for analyzing scientific reasoning in assessments” (Brown, Nagashima, Fu, Timms, & Wilson, 2010)

Claim	FV – Fully Valid PV – Partially Valid IV – Invalid, an incorrect link NL – No Link NC – No Claim
Evidence	SE – Sufficient Evidence GE – General Evidence IE – Insufficient Evidence UE – Unrelated evidence NE – No Evidence
Reasoning	CC – Connects to scientific principles and connects evidence to claim CP – Connects to principles CE – Connects evidence to claim NC – No connections made NR – No reasoning

APPENDIX F
POST INTERVIEW QUESTIONS

POST INTERVIEW QUESTIONS

1. When we do Claim Evidence Reasoning at the end of a lab, how confident are you that you can write a good claim, evidence, and reasoning?
2. What is the hardest part of doing a CER?
3. When someone else shares their CER how can you decide if their evidence is good?
4. When someone else shares their CER how can you decide if their reasoning is good?
5. When someone disagrees with what you know about science, how do you respond?
6. Which do you think is easier on a test, short answer or multiple choice? Why?
7. How confident are you in your ability to explain yourself on short answer questions?
8. Is there anything else you would like me to know?