WRITTEN ARGUMENTATION AND STUDENT LEARNING IN AN 8TH GRADE SCIENCE CLASSROOM

by

DAVID PAULI

(Under the Direction of David F. Jackson)

ABSTRACT

With the advent of the Next Generation Science Standards (NGSS), instruction in the science classroom needs to change (Lee, Quinn, & Valdés, 2013). The NGSS values argumentation in the science classroom (Lead States, 2013) by specifically naming argumentation as one of the eight scientific processes. Writing is also a valued instructional strategy (Keys, 1999). The problem is to define a practical, useful instructional strategy that blends argumentation and writing to promote student subject matter learning. Understanding the mechanisms by which writing and argumentation work together to promote student learning will enable researchers and teachers to enact writing and argumentation in the classroom efficiently. This study purposes to investigate the Answer, Cite, Organize, Respond, and Note (ACORN) as a new framework for written argumentation specifically designed to elicit metacognition. . In this mixed methods study, the three research questions were addressed as follows: Using the ACORN framework, to what extent can students produce a high scoring structured argument as measured by the Total Argument (TA) Rubric or the Holistic Argument (HA) Rubric? (Choi, Notebaert, Diaz, & Hand 2010) What metacognitive components do students engage in when using the ACORN framework? Do Total Argument scores and/or Holistic Argument scores on written arguments developed from the ACORN framework correlate with student subject matter learning? Eighth grade students (N=48) were taught a nine week unit on gravity, Newton's Laws, forces, speed, velocity, acceleration, and simple machines. During the course of regular instruction, students engaged in written argumentation and metacognition using the ACORN framework in five different writing tasks. Results showed that using the ACORN framework, a student can produce a high scoring, structured argument. Additionally, students who used the ACORN framework primarily engaged in metacognition through use of regulation of cognition. While there was only a weak correlation quantitatively between written argumentation ability and subject matter learning, qualitative evidence suggests otherwise. One important result to come from ACORN for argumentation was the ability of the students help distinguish between evidence and reasoning. Implications of this research include establishing the ACORN framework as tool for argumentation as well as directions for future research.

INDEX WORDS: Argumentation, Metacognition, Mixed methods, Middle school

WRITTEN ARGUMENTATION AND STUDENT LEARNING IN AN 8TH GRADE SCIENCE CLASSROOM

by

DAVID PAULI

Ed.S., Brenau University, 2000

MS., Georgia Institute of Technology, 1993

BA., Georgia Institute of Technology, 1992

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2017

WRITTEN ARGUMENTATION AND STUDENT LEARNING IN AN 8TH GRADE SCIENCE CLASSROOM

by

DAVID PAULI

Major Professor: David Jackson

Committee: Cory Buxton
Barbara Crawford
Carolyn Wallace

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Writing and argumentation are complex topics. Understanding the mechanisms by which the two work together to promote student learning will enable researchers and teachers to enact writing and argumentation in the classroom efficiently. The question then becomes what components of written argumentation promote student learning. Teachers and students can be intimidated by writing in the classroom (Street, 2003). Furthermore, teachers and students do not always understand the importance of argumentation in science (Duschl & Osborne, 2002; Duschl, Schweingruber, & Shouse, 2007). Finding a research based approach to combining these topics could possibly yield a powerful tool to promote student understanding of science topics.

Both writing (Keys, 1999) and argumentation (Duschl & Osborne, 2002) are forms of scientific discourse. Students must be given the opportunity to write in the classroom in order to hone the skills needed to promote leaning (Sampson, Enderle, Grooms, & Witt, 2013). In order for students to master argumentation, Duschl & Osborne (2002) suggest that science classrooms need to offer students the opportunity for scientific discourse. This discourse is commonplace in the scientific community, but not in the classroom (Duschl & Osborne, 2002). The practice of science educators must change to include the opportunities for argumentation and writing as a matter of scientific discourse as a part of regular classroom instruction to help promote student learning. Therefore, frameworks that provide options for written argumentation are a valuable tool for classroom instruction.

Searching for a practical instructional tool for the classroom is always a challenge. One tool is Restate Answer Cite and Explain (RACE) paragraphs. These paragraphs are used to help students write answers to questions using a structured format (Himmele, Himmele, and Potter, 2014). Even though RACE paragraphs are used as an instructional strategy in some science classrooms, no research exists that connect RACE paragraphs to argumentation or student learning. Furthermore, research does not exist connecting RACE paragraphs with metacognition, and there is no evidence that students engage in metacognition as a part of using RACE paragraphs. If RACE paragraphs are to be used in the classroom, then it would behoove educators to understand the advantages and disadvantages of this instructional strategy from a research perspective.

This study purposes to examine a new framework for written argumentation that combines the essential features of RACE paragraphs with reflection activities designed to elicit metacognition. I hope that by adding these metacognitive activities through writing and argumentation that a new, useful instructional tool based in research will be developed.

Definitions of Key Terms Adopted for this Study

Argumentation: “the coordination of evidence and theory to support or refute an explanatory conclusion” (Osborne, Erduran, & Simon, 2004, p. 995). Many definitions of argumentation exist (Erduran & Jiménez-Aleixandre, 2008), but this definition is practical in nature and is used to develop the conceptual framework for this study.

Metacognition: “the awareness and management of one’s own thought” (Kuhn & Dean, 2004, p. 270). This definition is a practical definition that allows for metacognition to be modeled and practiced in the classroom. Kuhn (1992) states that metacognition is the foundation for argument as thinking.

Scientific explanation: the end product, whether written or oral, of the coordination of evidence and theory in the context of science. While there is no one accepted definition of scientific explanation (McNeill, Lizotte, Krajcik, & Marx, 2006), for the purposes of this study, scientific explanation will be used to denote the product of argumentation.

Student learning of subject matter knowledge: the ability of a student to recall, apply, or synthesize knowledge specific to the science subject matter being taught. For the purposes of this study, the subject matter is traditional physical science (Newtonian mechanics) topics of gravity, Newton's Laws, forces, speed, velocity, acceleration, and simple machines

Written argumentation: a subset of argumentation that focuses on the writing process to coordinate evidence and theory.

Background of Study

This study is founded in my practical experience. During twenty-one years of teaching, I have always promoted writing as a means of learning, especially in science. Writing as a tool for science learning is well founded in the literature (Wallace, Hand, & Prain, 2004), but writing alone does not guarantee learning (Keys, 1999). One of the key components that links writing to learning is metacognition (Flower & Hayes, 1981; Mayher & Lester, 1983; Applebee, 1984, Langer & Applebee, 1987; Bereiter & Scardamalia, 2013).

Writing in all curriculum areas was a main initiative at my school. RACE paragraphs were readily accepted as a valued instructional strategy even though there is no research on RACE paragraphs. Furthermore, I was not seeing the correlation between RACE paragraphs and student subject matter learning as measured by standard classroom assessments. I initiated a study to examine the connection between writing RACE paragraphs and subject matter learning (Pauli, 2017). In this quasi-experimental study, the intervention group of students use RACE paragraphs to learn science content about nuclear processes while the control group used non-argument based recall and analysis questions to learn the same content. There was no significant difference in leaning between the two groups as measured by the unit test. Also using RACE paragraphs, students did not produce high quality argumentation. To further investigate the connections among writing, argumentation, and subject matter learning, the present study was designed.

Statement of the Problem

With the advent of the Next Generation Science Standards (NGSS), instruction in the science classroom needs to change (Lee, Quinn, & Valdés, 2013). The NGSS values argumentation in the science classroom as a matter of discourse (Lead States, 2013) by specifically naming argumentation as one of the eight scientific processes. Writing is also a valued instructional strategy (Keys, 1999). The problem is to define a practical, useful instructional strategy that blends argumentation and writing to promote student subject matter learning.

Purpose of the Study and Research Questions

The purpose of this study is to investigate the Answer, Cite, Organize, Respond, and Note (ACORN) framework as an instructional strategy. I have designed this framework to address the apparent lack of metacognition in RACE paragraphs. First, writings generated from the ACORN framework were mapped to argumentation by meeting the standards set by scientific and educational communities. Second, the experiences of the students using the ACORN framework were examined for the mechanisms of learning. Third, any relationships among the processes of using the ACORN framework, argumentation, and student subject matter learning were determined.

Research Question 1: Using the ACORN framework, to what extent can students produce a high scoring structured argument as measured by the Total Argument (TA) Rubric or the Holistic Argument (HA) Rubric? (Choi, Notebaert, Diaz, & Hand 2010) (Quantitative)

This question was designed to map the ACORN Framework to argumentation. The literature has already shown the benefits of argumentation through other frameworks. Establishing whether or not the ACORN framework can be seen as argumentation is the first step in understanding how the ACORN framework may help students learn.

Research Question 2: What metacognitive components do students engage in when using the ACORN framework? (Qualitative with Quantitative support)

This question was designed to examine metacognition through interviews, think alouds, and the Metacognitive Awareness Inventory (MAI) in addition to triangulating evidence of metacognition with TA and HA scores. This is the overarching mixed methods question driving this research. (Tashakkori & Cresswell, 2007). Notice that the question here takes on the "what" format as described by Hesse-Biber (2010) indicating a qualitative question.

Research Question 3: Do Total Argument scores and/or Holistic Argument scores on written arguments developed from the ACORN framework correlate with student subject matter learning? (Quantitative with Qualitative support)

This research question was designed to show if strong arguments from ACORN paragraphs can be correlated to student subject matter learning. Information from student interviews and the think alouds are used to explain the results.

Theoretical Framework

The theoretical framework through which this study was viewed is pragmatism. Pragmatism as theory has it roots with John Dewey (Denzen, 1996; Green, 2007) and also was a part of the history of the Chicago school (Denzen, 1996). Several of the general characteristics of pragmatism (Johnson & Onwuegbuzie, 2004) are valuable to me and fit well with the study at hand. Johnson and Onwuegbuzie (2004) list more general characteristics of pragmatism, but the ones presented below are the primary ones that informed this study.

Pragmatism holds that a reality exists outside of the learner yet that it is constructed by the learner (Johnson & Onwuegbuzie, 2004). Scientific inquiry of which argumentation is a part is based in constructivism (Abd-El-Khalick et al., 2004). At the same time the natural world exists outside of human constructs. The learner needs multiple avenues to make sense of that reality and written argumentation can be one of those avenues. Ontologically speaking, the nature of the world is that it exists outside an individual organism. At the same time, that outside world needs to be interpreted which is unique to an organism. The epistemology I associate with pragmatism is that we come to know by constructing meaning from the world around us. In line with pragmatism, meaning is constructed subjectively in an objective world. Part of Dewey's pragmatism is transactional realism that allows for both subjective and objective views of reality

(Biesta & Burbules, 2003). "What is constructed – over and over again – is the dynamic balance of organism and environment". (Biesta, and Burbules, 2003, p. 11).

Pragmatism rejects dualism (Johnson & Onwuegbuzie, 2004). Dualisms in a philosophical sense are mutually excusive ways of viewing the world. For example, free-will versus determinism (Johnson & Onwuegbuzie, 2004) is philosophical dualism that represents opposing viewpoints. Rarely does the practical world exist in dualisms. Argumentation shows this by giving students a variety of opportunities to reason through a scientific claim. By the nature of student writing, there are multiple ways to address the same issue with a variety of evidence and reasoning. Additionally, metacognition also cannot be characterized in binaries. Metacognition by definition is reflecting on one's thinking. Metacognition is unique to each individual and requires more than one way to examine this. Because one of the main reasons for the ACORN framework is to intentionally provide the opportunity for metacognition, I needed appropriate tools to understand student thinking. This provides some of the reasoning for addressing this study using a mixed methods approach. To provide an understanding of how metacognition plays a role in written argumentation, a quantitative and qualitative approach was necessary.

Pragmatism holds that the results of research are not final truths (Johnson & Onwuegbuzie, 2004). Similarly, science is constantly changing and adapting based on new evidence (Lederman, 2007). Argumentation is based in supporting a claim with warranted evidence. The process of argumentation in the classroom allows students to understand that the goal of science is to generate the best explanation given the evidence at hand. Students learn science content and overcome misconceptions similarly to how scientists explain the world. As a tenet of science, knowledge can change or be adapted based on new evidence. The axiology I associate with pragmatism is the ethical presentation of evidence in order to generate meaning. Pragmatism presents evidence in the goal of obtaining a truth, but recognizes that there is no absolute truth (Johnson & Onwuegbuzie, 2004).

Pragmatism holds that different views and perspectives have value (Johnson & Onwuegbuzie, 2004). For argumentation in the classroom to be effective, students need to assess different evidence to help explain science content (Duschl & Osborne, 2002; Osborne, Erduran, and Simon, 2004). In other words, different pieces of information need to be considered, their value weighed, and either accepted or dismissed as warranted evidence supporting explanations.

Pragmatism holds that results from research are considered "warranted evidence" (Johnson & Onwuegbuzie, 2004, p. 18) that lead to a better understanding of the bigger picture. This study was a follow-up to a previous study that investigated RACE paragraphs; moreover, this research was designed to fit in the larger picture of the argumentation literature.

Pragmatism "endorses a strong and practical empiricism as the path to determine what works" (Johnson & Onwuegbuzie, 2004, p. 18). In other words, evidence is critical to the research findings. In order to give the strongest account from this research study, a mixed methods design using multiple data sources in an attempt to triangulate results was employed.

Pragmatism is iterative (Johnson & Onwuegbuzie, 2004). The iterative nature of this study can be seen on two different levels. First, the initial quantitative results were used to help inform the design of semi-structured interview questions. This iteration is designed to gain practical and useful knowledge for the explanatory nature of this study. Second, the written framework under study is also iterative in nature since students organized and reflected on their work as will be discussed later.

Conceptual Framework

Argumentation has been linked to students learning either science subject matter knowledge or scientific processes (Sandoval & Millwood, 2007; Jiménez-Aleixandre & Erduran, 2007). Subject matter knowledge and scientific process are intimately intertwined (Erduran & Jiménez-Aleixandre, 2008) and each is complex in its own manner. Consider student learning of science whether it be subject matter or sciences processes as a structure containing a variety of supports as seen in Figure 1.1. The top of the structure, student learning, is held up in many different ways represented by different sides of the structure. Presumably the more supports, the more stable the structure.

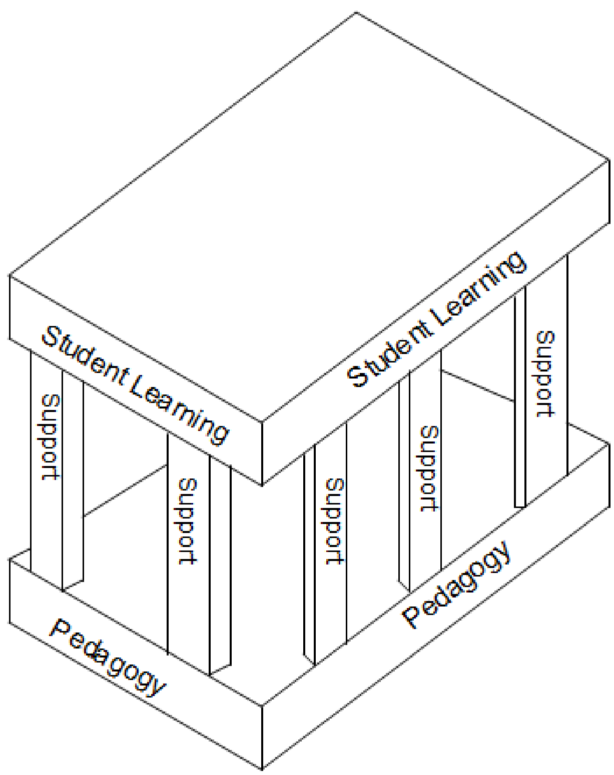


Figure 1.1. Support of Student Learning

In Figure 1.1, student learning is supported through pedagogy many different ways. Pragmatically, this makes sense as there is more than one way for a student to learn. Furthermore, just as any concrete structure can be built from a variety of materials, student learning can be built by any number of pedagogical supports. This conceptual framework looks at one side of the structure, one set of supports for student learning. Specifically, argumentation supporting student subject matter learning will be examined. This study does not imply that argumentation is the only support for student learning, but I claim that argumentation is a support for student learning.

Many different researchers have proposed a definition for argumentation (Sampson & Clark, 2008, Erduran & Jiménez-Aleixandre, 2008), but here I will use a definition proposed by Osborne, Erduran, & Simon (2004), "the coordination of evidence and theory to support or refute an explanatory conclusion" (p. 995). This definition is practical in nature and understanding how each word is used will result in a framework for argumentation. This framework shows how argumentation supports student subject matter learning of subject matter knowledge through scientific process and science content (See Figure 1.2).

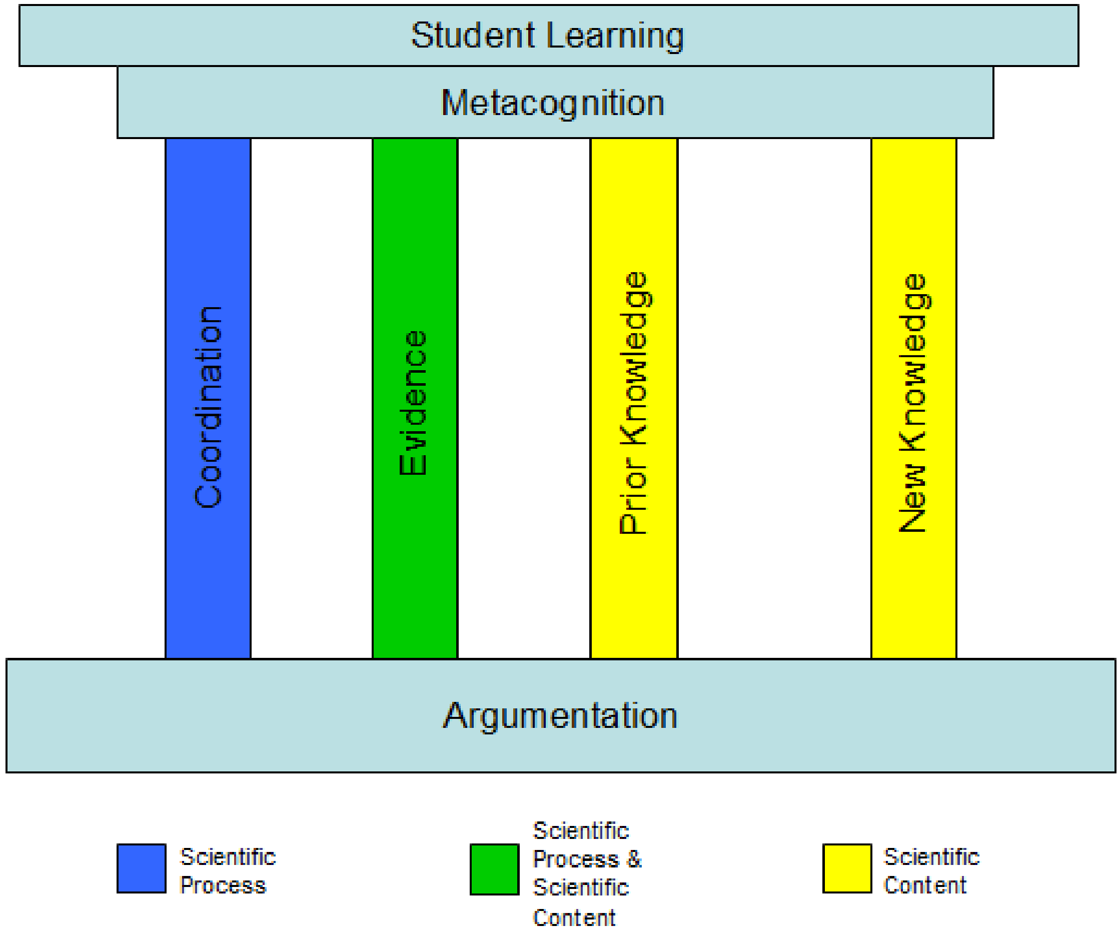


Figure 1.2. Aspects of Argumentation related to Student Learning

The use of the word coordination implies an active process involved in manipulating information. This manipulation is a scientific process necessary to participate in argumentation. The word evidence implies observable phenomena and data. These phenomena and data must be recognized by the student as valuable to the claim. Once this happens, phenomena and data qualify as evidence.

Evidence is partly science content, but also partly science process. What makes evidence both process and content is the understanding that evidence is different from new knowledge and prior knowledge. A student must be able to differentiate among prior knowledge, new knowledge, and evidence to understand that evidence is used to connect prior knowledge to new knowledge. The differentiation is scientific process at work. The word theory in this definition is not necessarily the classic definition theory in science. Theory here refers to scientific content knowledge that the student has already learned whether that content knowledge is a list of facts or a set of procedures. In other words, theory as applied here is prior knowledge which is important in thinking about science (von Aufschnaiter, Erduran, Osborne, & Simon, 2008). The explanatory conclusion is the new scientific knowledge that the student intends to learn. Using argumentation language, this is essentially the claim that is being made. Together, argumentation is an active process that links together prior knowledge with new knowledge. In other words, argumentation involves both scientific process and scientific content.

Argumentation involves metacognition (Erduran & Jiménez-Aleixandre, 2008; Garcia-Mila, & Andersen, 2007; Zohar, 2007), and metacognition supports student learning. Argumentation through coordination, evidence, prior knowledge, and new knowledge also supports student learning (Kuhn, 2000). Therefore, the processes associated with argumentation should also be intertwined with metacognition. Students must be taught how to think metacognitively (Garcia-Mila, & Andersen, 2007; Kuhn & Udell, 2003). When teaching argumentation, explicit instruction on thinking metacognitively also needs to be given.

Overview of Methods

The study took place in a suburban middle school with two classes of gifted 8th grade students taking a high school level physical science class. The nine-week unit of study covered traditional physical science (Newtonian mechanics) topics of gravity, Newton's Laws, forces, speed, velocity, acceleration, and simple machines. As a routine practice of instruction, students wrote using the ACORN argumentation framework.

Graphical representation

This study will take place in several parts over approximately a seven month period. Both qualitative and quantitative data will be collected and analyzed. Figure 1.3 was created to highlight the important components of the study (Creswell, 2014).

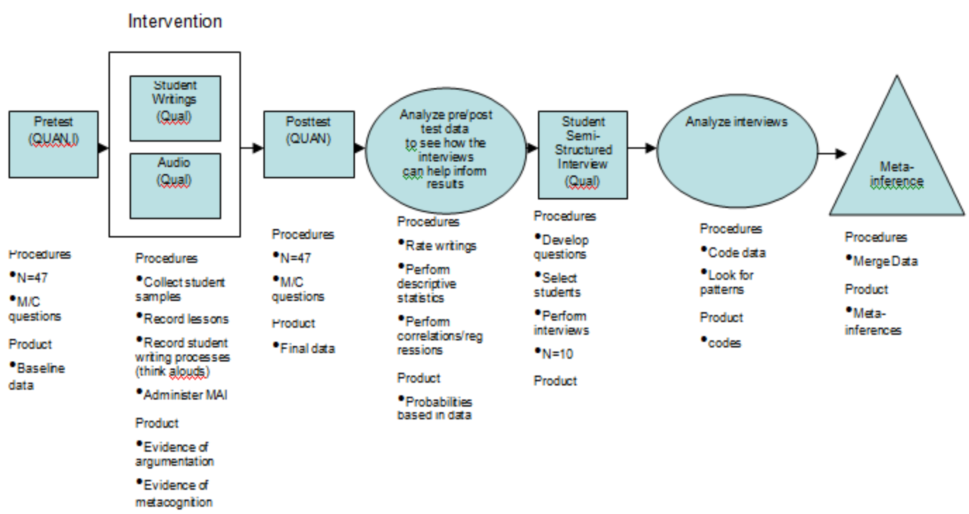


Figure 1.3. Graphical Representation of the Study

Significance of Study

This is study is significant because it will systematically examine a practical, generalizable framework designed to elicit written argumentation, metacognition, and student subject matter learning. While other frameworks for written argumentation exist, the ACORN framework attempts to incorporate the strongest aspects of those frameworks. By doing so, I hope to develop a sound instructional strategy that is easy to use and research based.

CHAPTER 2

REVIEW OF LITERATURE

Writing to Learn

Writing to Learn (WTL) is generally accepted as the use of informal writing or expressive writing to support learning (Kiefer, n.d.). In WTL, short writing activities are used to help students organize ideas and make connections to prior knowledge.

At the heart of WTL is the belief that writing promotes thinking which in turn promotes learning. Educators, researchers, and psychologists have worked over the last century to establish that link.

In his 1934 *Thought and Language,* Lev Vygotsky discussed the link between writing and thinking. He argues that writing requires greater abstraction, thought processes, and analytical action. He continues to define the difference between inner speech and written speech as "deliberate semantics – deliberate structuring of the web of meaning" (p. 182). He states that writing is a conscious act and cannot be fully comprehended until a child has matured. Vygotsky's work creates the link between writing and higher cognitive processes from a psychologist's point of view. Future researchers moved forward from this idea to establish a line of research that links writing and learning.

In *Writing as a Mode of Learning* (1977), Janet Emig describes a link between writing and learning. She begins by noting that writing is a unique process that develops only with "formal and systematic instruction" (p. 122). She also notes that writing is unique because the writer is originating a new record using graphical representation. She continues by making the distinction between talking and writing somewhat similar to Vygotsky's concepts of inner speech and written speech. She argues that talking and writing are two different processes whose differences are included in the Table 2.1 below:

Table 2.1. Writing Processes versus Talking Processes

Writing	Talking
A learned behavior	A natural behavior
An artificial process	Not an artificial process
The technological device	An organic device
Is slower	Is faster
stark, barren, and naked as a medium	Rich, luxuriant, and redundant
Provide its own context	Leans on the environment
Audience is absent	Audience is present
Results in a graphical representation	Does not result in a graphical representation
More responsible because a product is involved	Less responsible
Has a certain mystique	Treated mundanely
More readily a source of learning	Less readily a source of learning

The purpose of this argument is to lay the foundation for writing as an effective means of instruction. Where traditionally teachers lecture to students, the above list provides evidence that writing is superior to lecture.

Emig (1977) then laid out a definition of learning based on the works of Jerome Bruner and Jean Piaget that says we learn "by doing" (p. 124), we learn "by depiction in an image" (p. 124), and we learn "by restatement in words" (p.124). Emig argues that all three ways of learning are neatly encapsulated in the writing process. Next, she presents information about the workings of the brain citing several scientists in their research that both the right and left hemispheres of the brain are involved with writing. Then she follows up with the views of several psychologists including Vygotsky (1962, as cited in Emig), Luria (1971, as cited in Emig), Polanyi (1958, as cited in Emig), and Pirsig (1974, as cited in Emig) that support writing as learning.

The importance of Janet Emig's essay is that it is one of the first works that tries to link writing and learning together. While Emig does not supply any empirical evidence, it does lay the foundation in the literature of psychology that eventually led to empirical-based studies. Emig even suggests herself in the essay that this is the purpose by stating

> Yet I hope that the essay will start a crucial line of inquiry; for unless the losses to learners of not writing are compellingly described substantiated by experimental speculative research, writing itself as a central academic process may not long endure. (p. 127)

At the same time that Janet Emig's essay was presented, Linda Flower and John Hayes were working on their idea of a cognitive process theory based on writing. In *A Cognitive Process Theory of Writing*, they share their work of five years developing this theory. Their theory is used to describe the writing process as it relates to thinking. This theory contains four tenets which are described below. While the details of the model are too complex to be included in the scope of this paper, the importance of this theory as related to writing and learning will be highlighted.

Flower & Hayes (1981) described the first tenet as "The process of writing is best understood as a set of distinctive thinking processes which writers orchestrate or organize during the act of composing." (p. 366) The model of the day, and a model still touted in school today – the stage model – is the classical structure of writing that involves some sort of pre-writing activity, preparing a rough draft, and then revising. The authors argue that this model is too simplistic in that these processes are not linear. Writers are constantly planning, writing, and revising in no particular order. They continue to argue that the stage model does not allow for the complex interaction of the writer and the writing. The theory that they present claims that writing is a hierarchical process.

Flower & Hayes (1981) described the second tenet as “These processes have a hierarchical, highly embedded organization in which any given process can be embedded within any other.” (p. 366) The authors then go on to describe that their model by nature of its hierarchal structure allows for nonlinear thought processes. They argue that this nonlinear process involves powerful cognitive skills.

Flower & Hayes (1981) described the third tenet as “The act of composing itself is a goal-directed thinking process, guided by the writer's own growing network of goals.” (p. 366) Using this tenet, the authors compare the process of writing as seemingly disorganized yet directed in purpose. They make the distinction between process goals which are individual instructions on how to go about writing and content goals which are the overarching goals for the paper. An explanation of these goals is then used to reconcile how the writing process results in a final product.

Flower & Hayes (1981) described their fourth tenet as follows:

> Writers create their own goals in two key ways: by generating both high-level goals and supporting sub-goals which embody the writer's developing sense of purpose, and then, at times, by changing major goals or even establishing entirely new ones based on what has been learned in the act of writing. (p.366)

For this point the authors describe sub-goals and regenerating goals. Sub-goals are meant to help the writer accomplish more abstract goals, and regenerating goals are ones that have been changed because something has been learned through the writing process.

These four points provide one theory of how writing is directly related to thinking which in turn opens the door for later research to examine writing and learning. The first three key points suggests that there is a great amount of cognitive flexibility in the process of writing, and that this flexibility involves higher order processes place. One can argue that deep learning can take place only in the context of higher order thinking skills. The idea is we learn as we write. The fourth point, to be able to reevaluate goals based on what has already been written as a part of the writing process, implies that some learning must have taken place. One cannot reevaluate what has not been learned.

John Mayher and Nancy Lester (1983) attempt to redefine how student's learning takes place. The authors use science instruction as the vehicle to demonstrate how writing affects learning. They suppose that the traditional system of students learning bits of information by memorization is not as effective as students building a "useful, personal structure of knowledge". (p. 718) This would be accomplished through reflective, purposeful scaffolding of the curriculum and an active learning style, not so different from scientific inquiry. They propose that the writing process is both active and reflective meeting both of their criteria.

Arthur Applebee's *Writing and Reasoning* (1984) also establishes the connection between writing and thinking. According to the author, the four following factors connect writing to thinking:

> (a) the permanence of the written word, allowing the writer to rethink and revise over an extended period;
> (b) the explicitness required in writing, if meaning is to remain constant beyond the context in which it was originally written;
> (c) the resources provided by the conventional forms of discourse for organizing and thinking through new ideas or experiences and for explicating the relationships among them;
> (d) the active nature of writing, providing a medium for exploring implications entailed within otherwise unexamined assumptions (p.577)

Applebee is in line with the previous authors especially in relation to the active component of writing discussed in the fourth point and the connection to prior knowledge discussed in the third point. Having the time to reflect and relate information to prior knowledge is what writing does for thinking.

Applebee then shows that very little empirical research has been done to support writing and thinking. With regard to schooling he also suggests that writing activities are used limitedly and without much success. In research that he presents three years later, he provides the empirical research that he calls for and shows the link between writing and learning.

One foundational study that has connected science writing to student learning is Langer's and Applebee's *How Writing Shapes Thinking* (1987). In this empirical research, the authors set out to determine if writing supports thinking and how writing could be implemented in secondary classrooms. They argue that writing allows students to clarify their thinking by reflecting on the content. They further stated that even though writing is becoming more accepted as important to thinking, that writing instruction in the classroom is not improving. Specifically, their research looked at how different writing tasks impacted learning and how writing was use to support teacher's goals in the classroom.

This research task was a complex endeavor. It involved three years of research, several studies used to support the goals of the research. The participants included twenty-three teachers and five hundred and sixty-six secondary students all in the San Francisco Bay Area.

The authors conclude that "writing assists learning" (p. 135). They further delineate their findings by type of writing task. Below Table 2.2 summarizes the different types writing tasks and the consequences of using that task.

Table 2.2. Langer and Applebee's Explanation of Writing Tasks

Type of writing	Consequence
Short answer questions	• Information is often copied over with little learning going on. • Can promote good short-term recall of a lot of specific information
Analytic writing	• Ideas are dealt with in a more complex manner • Fewer ideas are dealt with • Promotes long-term learning
Summary writing and notetaking	• Related to the whole text • Little attention is paid to overall relationship

In Table 2.2, the different types of writing are linked to learning. Analytic writing, under which argumentation would fall, is associated with long-term learning. This is the link to WTL that gives written argumentation value as an instructional approach to promote student learning.

Langer and Applebee's (1987) work is valuable because it is one of the first studies that provides empirical evidence to connect writing to thinking. For a long period of time, writing as a means of learning was generally assumed by educators (Applebee, 1984). Providing a well-defined and extensive research study to present data and conclusions helped to establish the importance and permanence that writing should have in all disciplines.

In her work, "Revitalizing Instruction in Scientific Genres: Connecting Knowledge Production with Writing to Learn in Science", Carolyn Keys (1999) investigated different writing genres as they apply to learning science. She points out that there are several different viewpoints on the type of writing and how it promotes science learning. Keys (1999) makes the argument that students who write in traditional genres of science such as report, explanation, or experiment, do engage in learning. Keys (1999) states specifically, "Writing in the accepted scientific genres can provide opportunities for understanding the relationship of evidence to knowledge claims, and the tentative nature of the scientific enterprise" (p.119). Keys did not specifically study argumentation as a genre of writing even though the type of learning aligns with argumentation. Furthermore, many of the tenets of WTL such as organizing information, metacognition, and synthesizing new information overlap with argumentation.

A Brief History of Argumentation

Van Eemeren et al. (1996) report that argumentation was based in Greek logic established by Aristotle in the form of analytic or logic, dialectic or the theory of debate, and rhetoric or the verbal presentation of debate. The definitions of dialectic and rhetoric as applied to argumentation will change as modern theories are discussed. The assumption that Aristotle makes is that "all knowledge, insights, and opinions, insofar as they arise from rational thought, are based on existing knowledge, insights, and opinions" (Van Eemeren et al., 1996, p. 31). This idea from centuries ago lays the foundation of making scientific claims based on evidence.

The idea of argumentation as a form of instruction can also be traced to the 19th century where debates were held seen as sound pedagogy to prepare students for careers. It was also seen as a matter of citizenship. Debate primers were given to students in order to prepare them for careers that might use such skills (Van Eemeren, et al, 1996). From this point, argumentation evolves to the current understanding discussed at present.

Stephen Toulmin prompted the modern era of the study of argumentation in his book, *The Uses of Argument*, published first in 1958 and updated in 2003 (Erduran and Jimenez-Aleixandre, 2008). Toulmin (2003) makes reference to this in the preface of the updated printing of the book by suggesting that he did not intend to define a structure for argumentation. His original intent was to offer an opposing view to the idea that arguments needed to follow a strict and formal format similar to geometric proof.

Many science educators use Toulmin's original work or variations of it; therefore, understanding the basic structure that he presents is critical to the understanding of the field of argumentation. Toulmin (2003) proposes that argument can be broken down into six different categories: claims, data, warrants, backings, rebuttals, and qualifiers. Claims are the assertions or conclusions that have been proposed. The data are the facts of the situation. The warrants are the logical explanation of the data as it relates to the claim. While data may be specific to a case, warrants can be generalized in nature for any argument. Qualifiers are used to state how strongly the warrants justify the claim. Rebuttals provide the scenarios when the warrants must be set aside. Backings are the information that supports the warrants. Toulmin (2003) takes great care and provides examples of how each of these categories applies to an argument. Collectively, the structure of claims, data, warrants, backings, rebuttals, and qualifiers are referred to as Toulmin's Argument Pattern (TAP).

Conceptions of TAP

Bricker and Bell (2008) state that Toulmin's argumentation framework has been the most used framework in science education. Although TAP is widely used in science education, it does not come without criticism. One affordance of using TAP includes an easy structure to analyze (Van Eemeren, et. al. 1996). The process simply becomes a matter of taking the argument and identifying which parts of the argument fall into the categories of claims, data, warrants, backings, rebuttals, and qualifiers. The absence or presence of these components can then be analyzed qualitatively or quantitatively.

Another affordance of TAP is the "teachability." For students and teachers alike, a concrete structure to presenting an argument provides a place to begin to learn argumentation. Learning argumentation occurs best when related to prior knowledge (Von Aufschnaiter, Erduran, Osborne, & Simon, 2008). Since everyone has argued for something, even the smallest of children able to voice their concerns, being able to take an example and marry it to TAP is a good beginning place.

Yet a third affordance of TAP can be found in its commonality. Because it is used by many researchers, there exists a common language that promotes understanding among researchers. This is not to say different researchers do not have different takes on how to use TAP as a part of their research, but the basic definitions of claims, data, warrants, rebuttals and to a lesser extent backings and qualifiers gives a frame for a complex topic. Sampson and Clark (2008) offer a solid review of schemes used assess argumentation in the classroom, TAP being one of the six. The other five differ significantly from TAP, but Sampson and Clark (2008) acknowledge that TAP has "influenced many subsequent frameworks" (p. 465).

TAP also has some limitations. The first is that using Toulmin's structure does not take into account the content of the argument (Van Eemeren et.. al 1996, Toulmin, 2003, Sampson & Clark, 2008). Content is important when arguing (Bricker & Bell, 2008) and therefore causes a significant problem when evaluating the quality of students' arguments. Toulmin recognizes this himself when he states that each discipline needs to "judged by standards appropriate to that field." (p. 235). In other words, argument is dependent on the discipline.

The nature of this problem stems from the idea that TAP looks only for components or structure. It does not account for the quality of the statements in terms of scientific value (Sampson & Clark 2008). Van Eemeren et al. (1996) suggest that TAP is useful for the analysis but not evaluation of arguments. Consider the following example: The moon has different shapes (claim) because the shadow (data) is caused by the earth blocking the sun (warrant). While this argument contains three of Toulmin's components, the science behind this is wrong. To compensate for this, researchers add devices such as rubrics that measure the quality of the argument (Osborne, Simon, & Erduran, 2004; Choi, Notebaert, Diaz, & Hand 2010; Berland & McNeill, 2010)

Another limitation presented by TAP is the difficulty in establishing the difference between the different components of TAP. The difference between data and warrants seem to be of particular issue as noted by Van Eemeren et al. (1996) and Sampson and Clark (2008). Noting the 12-year span between these two articles suggests that this is not an easily solved problem. Consider the moon example above. The earth blocking the sun can be an explanation of how the shape of the moon is related to the shadow. On the other hand, the shadow caused by the earth blocking the sun could be considered as data. Teaching this distinction to students and teachers alike can provide a challenge.

Argumentation in Science

Currently, argumentation is a topic of interest in the science education community. Duschl and Osborne (2002) define argumentation as "the special case when the dialog addresses the coordination of evidence and theory to advance an explanation, a model, a prediction or an evaluation" (p. 55). They argue that this is a difficult task because it requires students to be aware of how to structure an argument, how to use information to make a convincing argument, and to be aware of the content that is being argued.

Osborne, Erduran, & Simon (2004) define argumentation as "the coordination of evidence and theory to support or refute an explanatory conclusion" (p. 995). They suggest that this view of argumentation supports both the learning of content and an understanding of the process of science. They argue that through the process of argumentation students can learn major concepts by being exposed to evidence and data. Once students engage in the process of argumentation to coordinate the information they are confronted with, students will learn the content (Osborne, Erduran, and Simon, 2004).

Erduran and Jiménez -Aleixandre (2008) state, "argumentation in scientific topics can be defined as the connection between claims and data through justifications or the evaluation of knowledge claims in light of evidence, either empirical or theoretical" (p. 13). They present a dual framework of argumentation. They suggest that argumentation has a social context as well as an internal context. Driver, Newton, and Osborne (2000) agree. Erduran and Jiménez-Aleixandre (2008) suggest that the internal process initiates thinking and that the social process calls for the negotiation of the information. The social context presents itself as a discourse among people and that discourse involves complex thinking skills. This leads to the idea that argumentation is a socially constructed act.

Sampson and Clark (2008) do not explicitly define argumentation, but do suggest a definition similar to Osborne, Erduran, and Simon by stating that argumentation is "the ability to generate a persuasive and convincing argument that coordinates evidence and theory to support or refute an explanation is an important component of the inquiry process" (p. 448). Here, scientific inquiry is related to argumentation and suggests that it is a critical component of inquiry. Most people accept that argumentation is a scientific discourse and in order to be enculturated into the scientific community, one must be explicitly taught argumentation (Osborne, Erduran, and Simon, 2004).

Each of the definitions are similar in that they each involve supporting a claim with evidence and being able to make the connection from the evidence to the claim. All of the above definitions also allow for argumentation to be oral or written. In a complex topic such as argumentation, though, it would be worthwhile to look at one venue for argumentation or the other since oral and written learning involve different processes. This is where the WTL research comes into play. The concepts behind WTL can be used as a lens to examine how written argumentation frameworks can support student learning.

Written Argumentation Frameworks

The following four written argumentation frameworks were critical in developing the new framework proposed for this study. Each of the frameworks will be briefly described.

The Science Writing Heuristic

The science writing heuristic (SWH) is a structured process used to make sense of data and concepts presented in a science laboratory activity (Wallace, Hand, & Prain, 2004). The process is guided through two different templates, one for the teacher and a corresponding one for the student. Both the teacher template and the student template can be found in Figure 2.1 (Nam, Choi, & Hand, 2011, p. 1114).

SWH template for teacher and student

Teacher template	*Student template*
Pre-Laboratory Activities: Teacher engages students to elicit pre-knowledge and gain understanding of the scientific context into which the aboratory is situated. Teacher may design pre-laboratory investigations such as brainstorming, developing questions about the topic, or expressing prior knowledge.	Questions: What are my questions?
Participation: Teacher encourages students to engage in an inquiry/laboratory investigation.	Test and Collect Data/Observation: What did I do? What did I see?
Negotiation I: Teacher guides students to think about the meaning of their data through journal writing.	Claims: What can I claim?
Negotiation II: Teacher encourages students to negotiate their understandings of the data with their peers. Students are encouraged to make knowledge claims to state explanations for their data.	Evidence: How do I know? Why am I making these claims?
Negotiation III: Teacher assists students to compare their ideas to textbook and on-line encyclopedia.	Reading: How do my ideas compare with others?
Negotiation IV: Teacher encourages students to communicate their current understandings of the investigation in a more polished form, i.e., writing a poem, letter or report, or creating a presentation or poster.	
Exploration: Teacher engages students to bring reflection to their understanding of the laboratory concepts.	Reflection: How have my ideas changed?

Figure 2.1 .Overview of the Science Writing Heuristic

In the pre-laboratory activity, students generate a question by connecting new information to prior knowledge. Wallace, Hand, & Prain (2004) suggest that this should be accomplished by a concept map, but in later iterations of the model, the design of the pre-lab activity is left to the discretion of the teacher. In the Participation part of the lab, students attempt to collect and organize data to answer the question or questions they generated in the pre-lab. The next four steps are negotiation processes to help the students make sense of the data they have collected and to make connections to important science concepts. In Negotiation Phase I, student use writing as a tool to make sense of their data. In Negotiation Phase II, students are beginning to form arguments through a discussion process with their peers. In Negotiation Phase III, the claims and evidence of the argument are compared to an authority, typically a written source such as a textbook. In Negotiation Phase IV, students produce a written document to present their argument. In the last step of the process, exploration, students are asked to reflect on their findings, writing, and experiences.

In this model, metacognition is evident in the Exploration phase. Reflection is evident in all four of the negotiation phases. Synthesizing information is evident in Negotiation Phases 1 and 2, as well as the Exploration phase.

Argument Driven Inquiry

Argument driven inquiry (ADI) is a structured process used to make sense of an inquiry based activity (Grooms, Enderle, & Sampson, 2015). Students use this framework by following eight steps to develop written argumentation. The steps are shown in Figure 2.2 (Grooms, Enderle, & Sampson, 2015, p. 47).

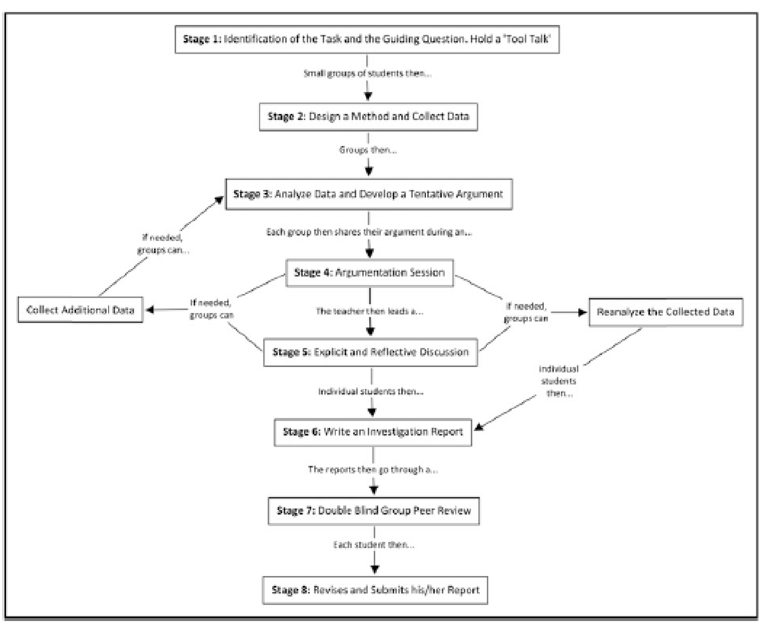


Figure 2.2. Argument Driven Inquiry

Grooms, Enderle, and Sampson (2015) describe an eight-stage process. In stage 1, students are identifying a question to be investigated through the inquiry process. In stage 2, students design the experiment, run the experiment, and collect data. At stage 3, the students develop an argument that would include a claim, evidence, and reasoning. Stage 4 is where students verbally make the argument to several of their peers. The peers evaluate these arguments and more data can be collected if necessary. Stage 5 is the security measure for the teacher to make sure that students are learning correct scientific concepts by having an explicit reflective discussion. Stage 6 is when students have the opportunity to produce the written argument which is labeled as an Individual Report. These reports are then reviewed by peers in a double-blind process during Stage 7. Stage 8 is the final revision by the original author of the Individual Report.

In this model, metacognition is evident in stages 5 and 8. Reflection is evident in stages 3 through 8. Synthesizing information is evident in stages 3, 4 and 6.

Claims, Evidence, Reasoning

Claims, Evidence, Reasoning (CER) is a non-structured model of written argumentation. In this framework, students are given the definitions for claims, evidence, and reasoning. While there are some variations in the literature for these definitions, McNeill and Martin (2011) offer these simplified definitions. They define a claim as a statement that answers a question. They define evidence as the data that supports a claim. They define reasoning as the explanation of how and why the evidence supports the claim using scientific ideas. Once these concepts are modeled, students write using this format and then receive feedback on their writing.

Metacognition and reflection are not directly evident in this framework. Synthesizing information is presumed as a part of the reasoning process.

RACE Framework

The proposed writing framework for this study was modified from the RACE framework. The RACE framework is a structured linear tool for students to answer constructed response questions (Himmele, Himmele, & Potter, 2014). The key to using the RACE framework is to understand the acronym stands as explained in Table 2.3.

Table 2.3. RACE Framework

Letter	Meaning
R	Restate the question by reflecting the wording from the original question
A	Answer the question using key words or phrases
C	Cite evidence from text or other appropriate source (lab data for example)
E	Explain how the cited evidence answers the question

Himmele, Himmele, and Potter, (2014) report that this framework has been used in the Conestoga Valley School District in Lancaster, Pennsylvania, with some success. This framework is used district wide from third grade to twelfth grade. As students progress to higher grades more complexity is added. For example, middle school students are expected to cite and explain more than one piece of evidence, and high school students are expected to correlate several sources to answer the question. The authors state that teachers have seen growth in students' written responses on standardized test.

Teaching students explicitly how to answer constructed responses is critical (Himmele, Himmele, & Potter, 2014). The RACE framework is straightforward, easy to memorize, and is good for a variety of students. The authors say that the teachers of the Conestoga Valley school district use this framework with all students and have seen benefits for ELL students and special education students. Special education students to gifted students have been taught effectively how to use the RACE framework resulting in greater clarity and structure in writing.

The challenge for most students in using this framework is the explanation of the cited information (Himmele, Himmele, & Potter, 2014). For the Explain step, a student must be able to make a relation between the cited information and the answer to the question. This cannot be a summary of what has been stated before and forces the student to integrate prior knowledge into answering the question (Himmele, Himmele, & Potter, 2014).

Metacognition

Effective science instruction must include the development of metacognitive skills, and these skills are necessary for higher levels of science (Schraw, Crippen, & Hartley, 2006). Metacognition has been generalized to the point where many teachers think that metacognition refers to any thinking strategy. Kuhn & Dean (2004) state that "We cannot effectively teach cognitive skills in the absence of very clear and precise understandings of what those skills are" (p. 269). Metacognition is a complex construct though and needs to be addressed as such (Baker & Cerro, 2000). Teachers are always looking for ways to teacher their students how to think, but often don't refer to the literature or understand how to the importance research (Kuhn & Dean, 2004; Baker & Cerro, 2000).

Some researchers suggest that argumentation involves metacognition (Erduran & Jiménez-Aleixandre, 2008; Garcia-Mila, & Andersen, 2007; Zohar, 2007). Erduran and Jiménez-Aleixandre (2008) state that argumentation supports "the access to the cognitive and metacognitive processes" (p. 5). Garcia-Mila and Andersen (2007) go even further suggesting that metacognition is needed to coordinate theory and evidence, an essential component of the definition of argumentation used here. Zohar (2007) indicates that argumentation necessarily involves metacognition since argumentation puts people in a situation where they are “immersed in the particular details of the cases they are considering and to ignore their deep logical structures.” (p.253) This forces people to enter into metacognition.

Definitions of Metacognition

Defining metacognition is a difficult task (Baker & Cerro, 2000). Kuhn and Dean (2004) define metacognition as “the awareness and management of one’s own thought” (Kuhn & Dean, 2004, p. 270). I have chosen this definition of metacognition since the work of Kuhn overlaps with metacognition and argumentation. The definition is practical in nature and gives a structure to examine the concept. As “awareness” and “management” imply, metacognition is generally divided into two categories: knowledge of cognition (awareness) and regulation of cognition (management) (Pintrich, Wolters, & Baxter, 2000; Schraw, 1998; Dean & Kuhn, 2004, Kuhn, 2000).

Knowledge of cognition is defined as what students know about their own cognition (Schraw, 1998). Knowledge of cognition can be subdivided into three domains, declarative, procedural, and conditional. Declarative is knowledge about oneself as a learner and what factors affect ones performance. Procedural is strategies, knowing how to do things, and includes categorizing and chunking information (Schraw, 1998). Kuhn (2000) further describes the knowing of procedural knowledge as either metastrategic or metatask. Metastrategic indicates knowing about the strategies, while metatask indicates knowing about the task to be accomplished. Conditional is when and why to use declarative and procedural knowledge (Schraw, 1998).

Regulation of cognition is a set of activities that one uses to control their learning and can also be subdivided into three domains: planning, monitoring, and evaluation. Planning involves the selection of strategies and the deciding how to apply them. Monitoring entails ones awareness of their progress with regular self testing. Evaluation is deciding the quality of products as well as the quality of ones learning (Schraw, 1998).

Assessing Metacognition

Schraw, Crippen, & Hartley (2006) note that metacognition is difficult to assess because the processes associated with metacognition are highly automated. Baker and Cerro (2000) suggest that observation of students performing authentic tasks is the simplest way to assess metacognition. These observations can be interpreted through observational checklists or by interview a student and having that student watch video recording and talk about what they are doing (Cerro & Baker, 2000).

One way to assess metacognition is through verbal reports which include interviews, questionnaires, and think aloud strategies (Baker & Cerro, 2000) While there are concerns with verbal reports, researcher generally accept that these are valid and reliable (Baker & Cerro, 2000). Keys (1999) used a think aloud protocol in her research with writing and the SWH. Based on the work of Ericsson and Simon (1980), Keys instructed students on a think aloud protocol, recorded students as they wrote, and then used those recordings as qualitative data. Upon analysis of the data, Keys was able to distinguish different examples metacognition, although these were not labeled as such. Keys (1999) states

> Students engaged in the following types of thinking while writing a laboratory report: generating hypotheses; generating evidence; evaluating hypotheses and evidence; generating meaning for patterns in the data; making general claims about the severity of the erosion problem (content space thinking); rehearsal of language choice for writing (movement from content space to discourse space); and rhetorical planning (discourse space thinking). (p. 683)

Evaluating hypotheses and evidence would fall under the evaluation or monitoring subcategories of metacognition. Rhetorical planning would fall under the procedural subcategory of metacognition. Using these think aloud strategies would be a valuable tool in gathering qualitative data.

The Metacognitive Awareness Inventory (MAI) is an inventory developed by Schraw and Dennison (1994) to measure the different components of metacognition (Baker & Cerro, 2000). This is a fifty-two-question student survey that differentiates between the main components of knowledge of cognition and regulation of cognition. Schraw and Dennison report that the MAI is reliable in measuring the main, but not the individual subcomponents such as the procedural metacognition or evaluation cognition. However, the MAI is easy to use in the classroom and are easy to administer (Pintrich, Wolters, & Baxter, 2000).

The inventory's questions are divided into two aspects of metacognition: knowledge of cognition and regulation of cognition (Schraw & Dennison, 1994). Table 2.4 shows which question fall into which category:

Table 2.4. Mapping of MAI Questions to Type of Cognition

Knowledge of Cognition	Regulation of Cognition
3,5,10,12,14,15,16,17,18,20,26,27,29,32,33,35,46	1,2,4,6,7,8,9,11,13,21,22,23,24,25,28,30,31,34,36,37,38,39,40,41,42,43,44,45,47,48,49,50,51,52

CHAPTER 3

RESEARCH DESIGN AND METHODOLOGY

Methodological Framework

Udo Kelle defines mixed methods research as

> Mixed methods means the combination of different qualitative and quantitative methods of data collection and data analysis in one empirical research project. This combination can serve for two different purposes: it can help to discover and to handle threats for validity arising from the use of qualitative or quantitative research by applying methods from the alternative methodological tradition and can thus ensure good scientific practice by enhancing the validity of methods and research findings. Or it can be used to gain a fuller picture and deeper understanding of the investigated phenomenon. (Johnson, Onwuegbuzie, & Turner, 2007, p. 120)

The goal of this research was to provide a better understanding of how metacognition can play a role in the ACORN framework. The instruments and data used are intended to support each other in examining this issue. More information on this will be provided when each of the proposed data sets are discussed.

Rationale

One of the main reasons for a mixed methods study was to examine the metacognition component. Research question 1 is quantitative since it is designed to map student writings into an argumentation format. By showing that students who use the ACORN framework have also demonstrated argumentation is the first step in correlating augmentation with metacognition through ACORN. In essence, I needed to verify that the student writings qualify as argumentation before I can look for a link among argumentation, metacognition and subject matter learning.

The second research question was designed to use qualitative information from the student writings, think alouds, and interviews with the quantitative data of student writings' TA and HA scores and MAI. Showing that students can produce high quality argumentation quantitatively is not enough to gain an understanding of how metacognition plays a role. A deeper examination of what the student was thinking is required. Triangulating quantitative and qualitative and quantitative results resulted in a richer understanding.

The third research question is primarily quantitative in nature, although a qualitative component can yield greater insight when dealing with a difficult to quantify concept of metacognition. Since eighth grade students are new to metacognition (Kuhn, 2011; Kellough & Kellough, 1999), think alouds, student interviews, and the MAI provided the best access to student metacognition. The quantitative HA & TA rubric data correlated to the FCI were compared with the qualitative data.

Purpose

The purpose for the mixed methods study was intended to be convergence as a result of triangulation (Green, 2007). Multiple data sets explained how metacognition and student subject matter learning are related to the ACORN framework. The first step was to establish ACORN as written argumentation through a quantitative measure. Second, metacognition was assessed quantitatively and qualitatively through students' think alouds, writings, and the argumentation process. Third, student subject matter learning was examined quantitatively and qualitatively. Some aspects of metacognition and argumentation as represented by the data triangulated to show convergence. As for student subject matter learning, the data diverged indicating a need for further research. Therefore, purpose of the study ultimately resulted in expansion (Green, 2007).

Type of mixed methods research design

The proposed research design method is an Explanatory Sequential Design (Creswell, 2014). In this design, quantitative data are collected and analyzed and then qualitative data is used help explain the results of the quantitative analysis (Creswell, 2014). The initial quantitative data analysis was centered on the pretest/post test scores as well as the TA and HA scores from various writings throughout the intervention. Qualitative data were collected from, student artifacts, think alouds, and semi-structured interviews. These data was used to help explain the quantitative results.

Intervention

The theoretical link between writing and learning is complicated and some researchers have suggested that metacognition is a key component (Wallace, Hand, & Prain, 2004). The purpose of this study is examine the Answer, Cite, Organize, Reason, Note (ACORN) framework as an alternative framework to promote argumentation and metacognition in science. The new framework is intended to combine the components of CER and RACE that promote an easy, generic framework for students to write with the metacognition associated with SWH and ADI. The intention of the acronym is to provide a structure that is easily understood and can be used by students and teachers for written products and the writing process. Table 4 explains the components of ACORN.

Table 3.1. Explanation of ACORN Acronym

Acronym	Concept	Explanation
A	Answer	Students answer the question or address a scientific claim given to them.
C	Cite	Students cite evidence related to the question or claim to support their answer. Evidence can come in the form of data from a lab, prior knowledge, or content knowledge.
O	Organize	Students perform an activity that will help them relate their cited evidence to reasoning. The nature of this activity is to elicit metacognition.
R	Respond	Students will write their reasoning developed from the previous step.
N	Note	Students will participate in an activity that allow them to reflect on how their reasoning is connected to the claim.

The introduction of Organize and Note steps are the critical ones that are missing in the RACE framework. One of the most difficult tasks in argumentation is to relate reasoning to the claims (Sampson & Clark, 2008). With the purposeful addition of the O and N steps, students will have an avenue to connect their reasoning to the claims engaging metacognition. The intent is to help students learn the material more effectively through the introduction of metacognition (Bangert-Drowns, Hurley, & Wilkinson, 2004) The O and N activities could possibly vary in nature from one writing activity to the next. For example, to help students organize their reasoning, students might use a graphic organizer, participate in think-pair-share activity, or complete a structured worksheet. To note how their reasoning connects to the claims, students could write in a science journal, peer review, or create a concept map. The idea is that O and N activities can be tailored to the task at hand, to the needs of the students, and the specific content being taught. This will prevent the writing process from becoming rote and stale.

Written argumentation frameworks have been well documented to promote student learning (Sampson & Gleim, 2009, Wallace, Hand & Prain, 2004; McNeil; McNeill, Lizotte, Krajcik, & Marx, 2006, Himmele, Himmele, & Potter, 2014). Table 3.2 compares ACORN to established frameworks in order show how the benefits of each are incorporated into ACORN.

Table 3.2. Criteria Comparison of Written Argumentation Frameworks

Criteria	**SWH**	**ADI**	**CER**	**RACE**	**ACORN**
Supports argumentation	Yes	Yes	Yes	Yes	Yes
Used to organize writing	Yes	Yes	Yes	Yes	Yes
Acronym is process driven	No	No	No	Yes	Yes
Used for writing outside of inquiry activities	No	No	Yes	Yes	Yes
Intentional Metacognition Activities present	Yes	Yes	No	No	Yes
Easily used for large and small classroom activities	No	No	Yes	Yes	Yes

Each of the criteria in Table 3.2 was selected for the practical application of written argumentation in the classroom. From a pragmatic perspective, the goal for this framework is based in argumentation and writing research, to be easily used in the classroom, and to be flexible enough to address a wide variety of practical needs. First, the ACORN framework needs to be based in argumentation and writing. Each of the frameworks that ACORN is drawn from were specifically developed with writing and presenting a scientific explanation (Sampson & Gleim, 2009; Wallace, Hand & Prain, 2004; McNeill, Lizotte, Krajcik, & Marx, 2006; Himmele, Himmele, & Potter, 2014). This is important because this establishes a solid foundation from the literature for ACORN. Second, the acronym is process driven meaning that each letter of the acronym is an action for a student perform. This helps to guide the student and the teacher through the writing and argumentation process during classroom instruction. Third, SWH and ADI were developed specifically for inquiry-based activities in the science classroom (Sampson & Gleim, 2009, Wallace, Hand & Prain, 2004). This limits their use. Practically speaking, time in the classroom cannot always be allotted for inquiry activities. ACORN is designed to be used for any application where a scientific explanation would aid in student subject matter learning. For example, ACORN could be used as an assessment tool, or a summary for a demonstration. Fourth, metacognition has been repeatly been shown to aid in student learning (Kuhn, 1992). ACORN specifically provides an avenue for metacognition.

Data Collection

Stage 1 consisted of research approval by all stakeholders. Permission to conduct the study was obtained from the principal of the school. In this district, since the research is being conducted by a teacher at his school, permission from the principal was sufficient and no additional approval from the district was necessary. I then requested approval from the university Institutional Review Board (IRB). Once the approval was granted, student and parent permission to participate in the study was issued and collected.

Stage 2 consisted of data collection and intervention. In order to address research question 3, students were given a pretest over the unit content. Students then engaged in the intervention process, the ACORN framework. During the intervention, student artifacts were collected and scored and student think alouds were recorded and transcribed to address research questions 1, 2, and 3. Additionally, the MAI was administered as a pre/post test to address research question 2. After students had the opportunity to engage in the intervention, students took a posttest to address research question number 3. Student semi-structured interviews were then conducted for the purpose of eliciting student experiences about the intervention, the writing process, and student metacognition to address research question 2.

Stage 3 consisted of data analysis. More detail will be given regarding this stage in the Data Analysis section.

Timeline

Stage 1, approval by all stakeholders, occurred during December 2016 and January 2017. Permission to conduct the study was obtained from the principal of the school in mid December 2016. IRB approval was requested and received during the first two weeks of January 2017. Once the approval was granted, student and parent permission to participate in the study took place in January 2017. Stage 2, data collection and intervention, will began at the end of January, 2017. The unit that was taught lasted approximately nine weeks, concluding in at the end of March, 2017. Stage 3, data analysis, occurred in two parts. The first part, the quantitative analysis, occurred immediately after the post test. Initial results were used to help inform the semi-structured interview questions (Creswell, 2014). The second part of the data analysis, the qualitative analysis of the interviews, think alouds and student artifacts, took place between March 2017 and August 2017.

Data Analysis

Data management

All data was collected and stored by me. Once the students who will participate in the study were identified through signed consent form, I assigned the students a research number that the students placed on all of the collected artifacts in lieu of their name. I also had a master list that linked the research number to student names. No other individuals connected to this study had access to this master list. Pseudonyms were used in reporting results for all interviews and think alouds.

Data analysis

Pretest/posttest

The instrument used for the pre/post test was the Force Concept Inventory (Hestenes, Wells, & Swackhamer, 1992). This inventory is a research-based test designed to measure student leaning in the areas of Newtonian Physics, gravity, forces and kinematics (Hestenes, Wells, & Swackhamer, 1992). The validity and reliability of this instrument has been well documented (Hestenes, Wells, & Swackhamer, 1992). Table 3.3 shows the alignment of the FCI to the intended topics of instruction. The only intended topic of instruction not found on the FCI is simple machines.

Table 3.3. Unit Comparison to FCI

Intended Instruction	FCI
Gravity	Gravity
Speed, Velocity, and Acceleration	Kinematics
Newton's Laws	Newton's Laws
Simple Machines	Not present

The pre/post test scores on the FCI were used for two purposes. First, the pre/post test scores were used to see if student subject matter learning took place. This was assessed using a paired t-test. The purpose for this analysis was to support research question number two. In order to describe the overall experience of using the ACORN framework, knowing if student subject matter learning has taken place helped to guide the interview process. The purpose of question number two was to determine how metacognition plays a role in argumentation. Second, the posttest scores on the FCI were correlated to the average TA for student writings and the score on the FCI will be correlated to average HA scores for the student writings. A linear regression was performed separately for each rubric score to determine whether or not there was a relationship between TA score and the FCI post test score and the HA score and the FCI post test score. This addressed research question number three.

The FCI is a secure document and the questions on the test cannot be shared (Hestenes, Wells, & Swackhamer, 1992). The developers of the FCI require training and agreements for the secure use of this document as a part of the integrity of the test (Hestenes, Wells, &

Swackhamer, 1992). The test is a well established instrument (Hestenes, Wells, & Swackhamer, 1992) that meets the needs of this study.

Student Writings

Kuhn and Crowell (2011) demonstrate that argumentation can be accurately assessed outside of Toulmin's pattern. Most researchers, however, use TAP in conjunction with rubrics to evaluate the quality of argumentation. For my study, I chose to modify two rubrics used by Choi, Notebaert, Diaz, and Hand (2010). While these two rubrics are primarily for assessment purposes, the rubrics can also inform instruction. Both rubrics, the Total Argument Score and the Holistic Argument Score rely heavily on Toulmin's structure in that relationships between claims and evidence (data) are examined. The issues of the quality are addressed in the Total Argument Score by specifically looking at the quality of the argument.

Students produced five written assignments using the ACORN framework. These rubrics were developed by Choi, Notebaert, Diaz, and Hand (2010) to rate written argumentation. One subscale on the TA rubric, the scoring matrix for reflection, was modified to assess specifically how well the student uses the explanation to connect the evidence to the claim to fit the purposes of this study. These rubrics were used because they provide two different perspectives on written argumentation and are generic enough to be used for different frameworks. Additionally, using pre-established rubrics helped link the current student to previous findings in argumentation. For a detailed explanation of the rubrics, see Choi, Notebaerk, Diaz and Hand (2010). Both rubrics can be found in the Sample Instruments/Protocols section.

Student interviews

After the posttest analysis has been completed, the results will be used to guide interview questions for a semi-structured interview (Creswell, 2014). Once the interviews were conducted, they will be transcribed. Using the constant comparative analysis (CCA) method (Glaser, 1965), the interview data was then coded, continuously comparing incidents to develop themes. These themes were then used to develop ideas to help address research question two.

Think alouds

During each writing instance, students will engage in think alouds, a verbal reporting protocol. During this reporting I hope to find instances of metacognition.
While students are engaging with the ACORN framework while developing their written argumentation, they will be recording all of their thoughts as they write. This protocol was modeled and practiced during all writing instances.

Ericsson and Simon (1993) report that this form of verbal report does not hinder a student's performance on a task. They report, "In general, instructions to think aloud do not interfere with subjects' thought processes, although they may slow them down somewhat if the subject is using a non-verbal code in his thinking" (p. 10). They also suggest that this type of reporting is reliable because it captures a student's thought as it is occurring. I am interested in whether or not students engage in metacognition, and this protocol allows me a glimpse of what is going on in the mind of the student.

Data Integration

Data was integrated at different points during the research. Student writing samples, initially in a qualitative form were quantified by scoring them with the TA and HA rubrics. This data transformation assisted in answering research questions one, two, and three. These data can were then compared to the FCI posttest scores to address research questions two and three. These data were analyzed, and further integration occurred with the development of the student interview questions. Student think alouds were examined both qualitatively and quantitatively in order to coborate information from all of the other data sets. The results from the analysis of the interviews and think alouds were used to triangulate data in the case of convergence (Green, 2007) and for expansion in the case of divergence (Green, 2007).

Participant Selection

The site at which the research took place was at a middle school in a suburban county in the Atlanta, Georgia metropolitan area. The student body at the school was comprised of 69% Caucasian students, 13% African-American students, 13% Latino students, 4% Asian students, and 3% multiracial students. Those eligible for free and reduced lunch made up 20% of the population. The school had an enrollment of 1658 students in grades six through eight. Students were grouped by teams according to the middle school concept in grades 6 and 7. Regular education and special education students were distributed randomly and heterogeneously among teams of teachers. Gifted students were grouped together in a separate team.

Students in grade 8 are scheduled according to the departmental/high school model. Gifted students are practically teamed, though, as most students in the gifted program take all of their four core academic subjects – math, language arts, science, and social studies from a total of 6 teachers.

While the school was selected out of convenience, additional factors make this site a valuable selection for research. This school has maintained a consistent and pervasive focus on writing and reading literacy. For the past eleven years, the teachers at the school have been receiving professional development on how to implement writing and reading strategies in order for students to be successful. As a result, the school has scored consistently in the top three schools in the state of Georgia on the eighth grade writing tests as well as on state standardized testing.

Quality Assurance

Quality of the data in the study will be addressed at two different levels. For the first level of quality control, each instance of data collection and/or analysis was assessed for quality according to that tradition. O'Cathain (2010) refers to this as the individual components approach. First, the number of participants is 48. This is an acceptable number to run the descriptive statistics, the t-tests, and the linear regressions. The FCI is a reliable and valid test that been used often in the science education research field (Hestenes, Wells, & Swackhamer, 1992). The rubrics used to score the student writing samples have been developed and used in a previous argumentation study (Choi, Notebaert, Diaz, and Hand 2010). The CCA method of the student interviews and think alouds follow a well-established framework for qualitative analysis.

The second level of quality addresses the quality of the overall study. O'Cathain (2010) suggests a framework for assessing the quality of mixed methods research. This framework is divided into eight different domains. The first three domains are appropriate to consider during the proposal stage of the study: planning quality, design quality, and data quality. Under planning quality, a thorough literature review was done, the rationale for the study must be transparent, deplaning of the study must be transparent, and the study must be feasible (O'Cathain, 2010). Under design quality, the design must be transparent and suitable for the research questions, the design must be strong, and the design must be rigorous (O'Cathain, 2010). Under data quality, the study must show data transparency, data, rigor, appropriate sampling, and appropriate and rigorous analysis (O'Cathain, 2010).

Ethical Considerations

The research took place at the school at which I taught, and student participants were drawn from eighth grade physical science classes. The classes were selected because they were taught by me. The benefit of this was easy access and an established rapport with my students. The liabilities of this were that familiarity between research and student may affect students' behaviors and responses.

I have taught middle school science for 18 years, and for that entire time I have promoted writing as a means to learn science. I believe the main mechanism for this subject matter learning is the process of students taking time to think and write a claim supported by evidence. I currently hold an Educational Specialist degree in the area of Math and Science Education. The benefits of being a teacher are having an understanding of how middle school students think and write. Also, working with middle school students for several years will benefit me in the interview process. The drawback to this research is the difficulty of being able to look for patterns or outcomes beyond what I have experienced as a teacher.

I am a White male teacher. There are no direct benefits for this. The concerns are that I will be propagating the myth of the White male scientist and that some students regardless of their own race or gender may stereotype me.

CHAPTER 4

RESULTS

The data for this research comprise several different quantitative and qualitative data sets. Each set will be described and analyzed separately as appropriate for an Explanatory Sequential Design (Cresswell, 2014).

Of the fifty-eight students in the two classes, forty eight students had completed data sets including Parental Consent, Assent forms, MAI pretest, MAI Post Test, and at least four out of the five sets of writings with an ACORN graphic organizer. Of the ten other remaining students, one did not give assent, three did not complete the FCI pre-test, five did not have at least four writings, one did not complete the FCI post test and one refused to take the MAI post test. All quantitative data analysis was used with N=48.

Quantitative Data Sets

For all quantitative variables, frequency histograms were generated and examined to ensure that the distribution of these variables appears to conform tolerably well to a normal distribution (neither highly skewed in either direction nor distinctly bimodal), thus satisfying that assumption of the conventional descriptive and inferential statistics used here. For all paired t-tests, $\alpha=0.05$.

Force Concept Inventory

The FCI was administered as a pretest before any content for this unit was taught. The FCI was then administered as a post test the day after the last writing. All guidelines for administering the FCI (Hestenes, Wells, & Swackhamer, 1992) were strictly followed. Students recorded answers on a multiple-choice scoring document and then were scored used All In Learning, a web based program that uses a document camera to score the multiple choice answers as well as aggregate the data. All FCI scores are represented as a percentage.

The FCI pretest data had an average of 21.73, a median of 20, a range of 33.3 and a standard deviation of 7.8. Figure 4.1 shows the histogram for the FCI pretest data.

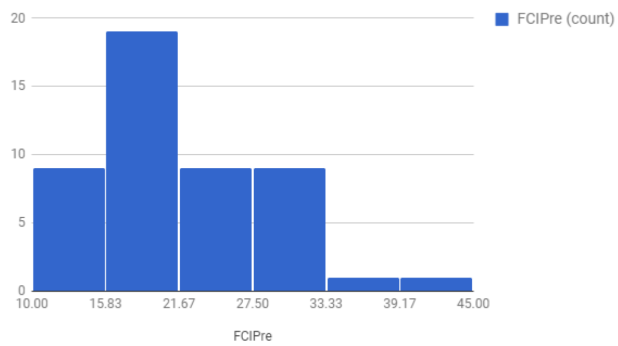


Figure 4.1. Histogram of FCI Pretest Data

The FCI post test data had an average of 29.16, a median of 30, a range of 60 and a standard deviation of 10.6. Figure 4.2 shows the histogram for the FCI post test data.

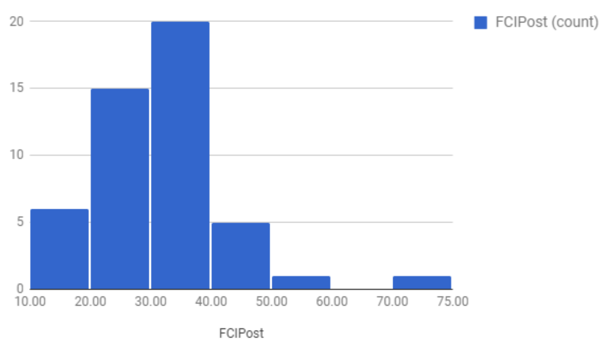


Figure 4.2. Histogram of FCI Post Test Data

A paired t-test was run between the FCI pretest scores and the FCI posttest scores. The difference of the means was found to be significant with p=0.00002. This indicates that students learned content during the time of the intervention.

Metacognition Awareness Instrument

The MAI was administered as a pretest before the first writing and as a posttest after the fifth writing. Both pre and post tests were administered on the computer using SurveyMonkey.com. This website was used in order to facilitate the collection of data in an electronic form. The original MAI was administered with fifty-two questions with participants marking their answer on a 10 mm line with one end of the line indicating false and the opposite end of the line indicating true. Responses were then measured to the closest mm giving a degree to the response (Schraw & Dennison, 1994). In the online version students moved a slider to indicate true or false on a scale of 0 to 100 giving the same resolution to the scale as the original MAI. Later versions of the MAI use a binary indicator for each question, but I agree with Schraw and Dennison (1994) that the ability of the respondent to give a varied response gives insight and information in assessing metacognitive awareness.

The MAI pre test data for the knowledge of cognition questions had an average of 78.0, a median of 68.2, a range of 53.6 and a standard deviation of 12.8. Figure 4.3 shows the histogram for the knowledge of cognition pretest questions.

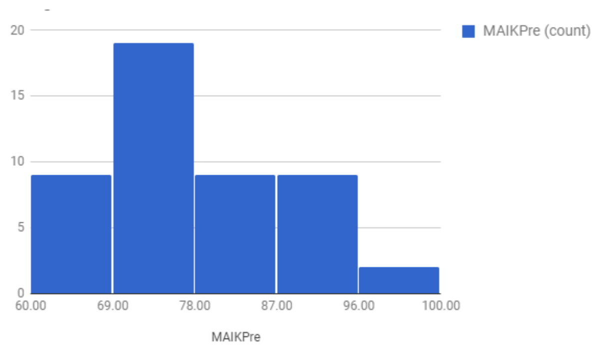


Figure 4.3. Histogram of Knowledge of Cognition Pretest Questions

The MAI Knowledge of Cognition post test data had an average of 79.8, a median of 78.9, a range of 60 and a standard deviation of 11.0. Figure 4.4 shows the histogram of the knowledge of cognition posttest questions.

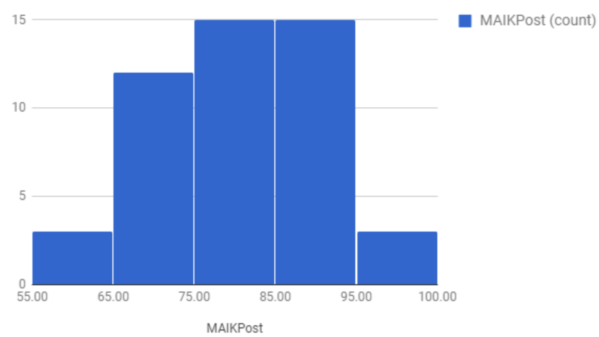


Figure 4.4. Histogram of Knowledge of Cognition Posttest Questions

A paired t-test was run between the MAI Knowledge of Cognition pre test scores and the MAI Knowledge of Cognition post test scores. The difference of the means was found to be not significant with $p=0.218$ and $\alpha=0.05$. This indicates that students did not have a significant change in their knowledge of cognition during the time of the intervention.

The MAI pre test data for the knowledge of regulation questions had an average of 66.6, a median of 68.21, a range of 53.59 and a standard deviation of 12.75. Figure 4.5 shows the histogram for the knowledge of regulation pretest questions.

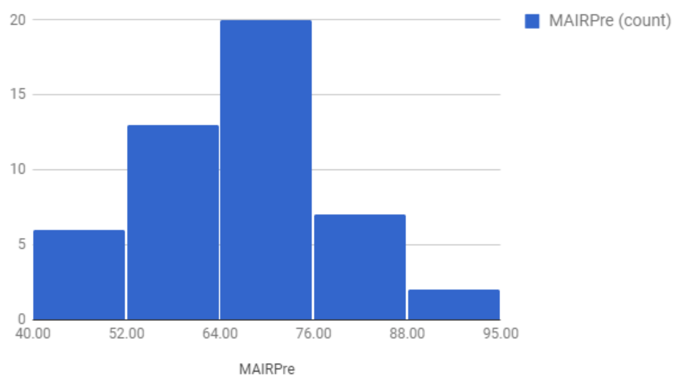


Figure 4.5. Histogram of Knowledge of Regulation Pretest Questions

The MAI Knowledge of regulation post test data had an average of 70.8, a median of 78.9, a range of 60 and a standard deviation of 11.0 Figure 4.6 shows the histogram for the knowledge of regulation posttest questions.

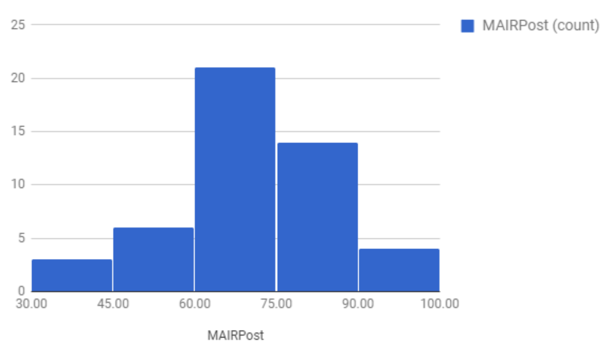


Figure 4.6. Histogram of Knowledge of Cognition Post Test Questions

A paired t-test was run between the MAI Knowledge of Regulation pre test scores and the MAI Knowledge of Regulation post test scores. The difference of the means was found to be significant with $p=0.006$ and $\alpha=0.05$. This indicates that students did have a significant change in their knowledge of regulation during the time of the intervention.

Student Writing Samples

The participants used the ACORN framework to produce written argumentation during five different instances. At the beginning of the intervention, the students were familiar with the concepts of claims, evidence, and reasoning, but they had not been introduced to the ACORN framework. The purpose of iterative instances of using ACORN was to allow the participants to learn the framework and to become comfortable using the framework in the context of argumentation. A discussion of each writing instance follows.

Gravity Writing

The Gravity Writing was the participants introduction to the ACORN framework as well as the introduction to the think aloud protocol. The students had just completed several days of instruction on gravity focusing on in particular gravity as a force, the gravity equation, how distance and mass affect gravity, and acceleration due to gravity.

For this writing assignment, the participants were randomly placed in groups of three or four. The researcher then modeled the use of the ACORN Framework as well as the think aloud protocol. Students were given a handout that described the ACORN framework and a graphic organizer (See Appendix). Participants then addressed the following problem using the ACORN framework with the graphic organizer.

> Mat and Nic were arguing about gravity. Mat said that gravity is not a constant on earth while Nic said that gravity was a constant on earth. Who is right? Using the ACORN framework, produce a written response.

Participants were allowed to discuss the problem and the process of writing as a group, but each individual completed their own writing. Furthermore, it was not necessary for student to agree as a group to complete the task.

Individual responses were first rated using the TA rubric. During this rating, each paper was marked showing instances of claims, evidence, and reasoning. Each instance was then compared the TA rubric and rated accordingly. For the Gravity writing, the average TA score was 11.40, the median TA score was 11, the range of the TA scores was 10 and the standard deviation of the TA scores was 2.4. No outliers were present in the data. Figure 4.7 shows the histogram for the gravity writing TA scores.

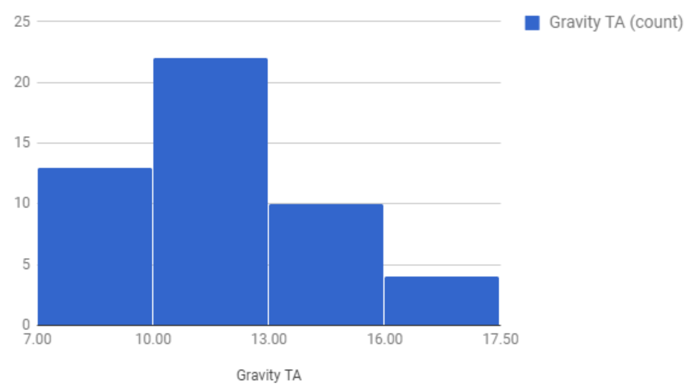


Figure 4.7. Histogram of Gravity Writing TA Scores

Individual responses were then rated using the HA rubric. During this rating, the writing was read holistically to gauge the overall effectiveness in argumentation. For the Gravity writing, the average HA score was 4.48, the median HA score was 4, the range of the HA scores was 5 and the standard deviation of the HA scores was 1.3. No outliers were present in the data. Figure 4.8 shows the histogram for the gravity writing HA scores.

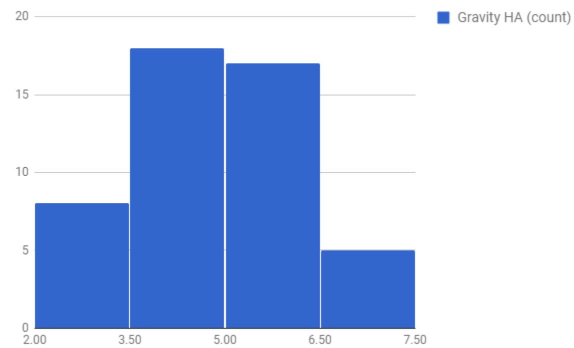


Figure 4.8. Histogram of Gravity Writing HA Scores

Newton's Three Laws Writing

The Newton's Three Laws Writing was the participants' second experience with the ACORN framework. The students had just completed several days of instruction on each of Newton's laws of motion and how they affect large body motion. For this writing assignment, the participants worked individually to answer the writing prompt. The researcher again modeled the use of the ACORN Framework as well as the think aloud protocol. Students used the same graphic organizer as the previous ACORN writing assignment. Participants then addressed the following problem using the ACORN framework with the graphic organizer.

> Which of Newtons Three Laws best represents a trebuchet? Using the ACORN framework, produce a written response.

Participants were not allowed to discuss the problem and the process of writing as a group. Each student was at his or her own computer and completed the writing using a word processor as well as a microphone headset to capture the think aloud process. Each student had the freedom to choose their own law to argue.

Individual responses were first rated using the TA rubric. During this rating, each paper was marked showing instances of claims, evidence, and reasoning. Each instance was then compared the TA rubric and rated accordingly. For the Newton's Law writing, the average TA score was 15.38, the median TA score was 15.5, the range of the TA scores was 10 and the standard deviation of the TA scores was 2.41. Figure 4.9 shows the histogram for the Newton's three laws writing TA scores.

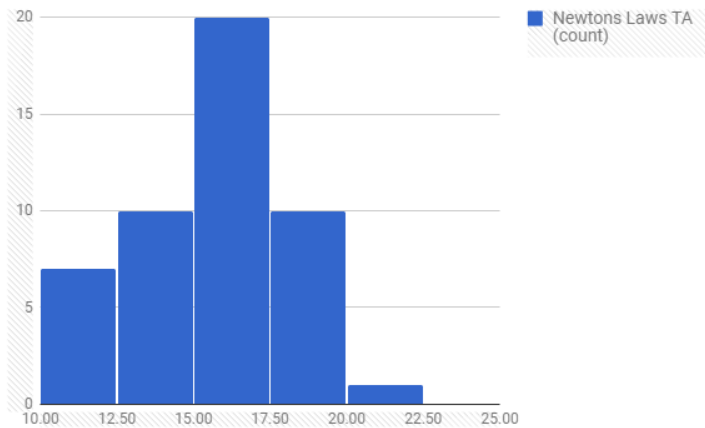


Figure 4.9. Histogram of Newton's Laws TA Scores

Individual responses were then rated using the HA rubric. During this rating, the writing was read holistically to gauge the overall effectiveness in argumentation. For the Newton's Law writing, the average HA score was 6.48, the median HA score was 7 the range of the HA scores was 4 and the standard deviation of the HA scores was 1.3. Figure 4.10 shows the histogram for the Newton's three laws writing HA scores.

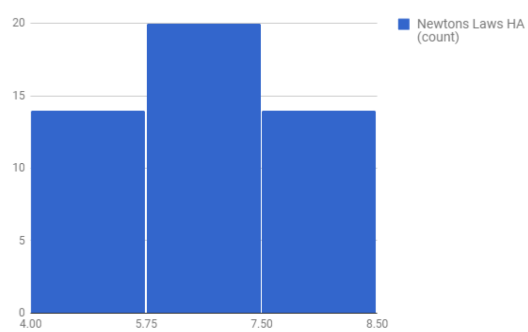
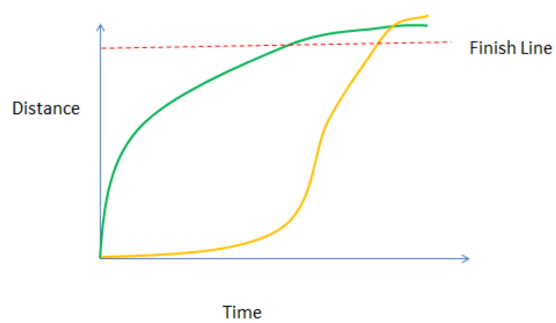


Figure 4.10. Histogram of Newton's Laws HA Scores

Speed, Velocity, and Acceleration Writing

The Speed, Velocity and Acceleration (SVA) Writing was the participants' third experience with the ACORN framework. The students had just completed several days of instruction on each on the definition and formulas for calculating speed, velocity, and acceleration. Specific attention was given to the participants on how to read, interpret, and draw time/distance graphs. For this writing assignment, the participants worked individually to answer the writing prompt. The researcher again modeled the use of the ACORN Framework as well as the think aloud protocol. Students used the same graphic organizer as the previous ACORN writing assignment. Participants then addressed the following problem using the ACORN framework with the graphic organizer.

> Using the following graph, decide who won the race and why. Use the ACORN framework to provide a written response.



Participants were not allowed to discuss the problem or the process of writing as a group. Each student was at his or her own computer and completed the writing using a word processor as well as a microphone headset to capture the think aloud process. Each student had the freedom to choose their own runner to argue although all 48 students choose the correct answer.

Individual responses were first rated using the TA rubric. During this rating, each paper was marked showing instances of claims, evidence, and reasoning. Each instance was then compared the TA rubric and rated accordingly. For the SVSA writing, the average TA score was 17.05, the median TA score was 17.5, the range of the TA scores was 5 and the standard deviation of the TA scores was 1.43. Figure 4.11 shows the histogram for the SVA writing TA scores.

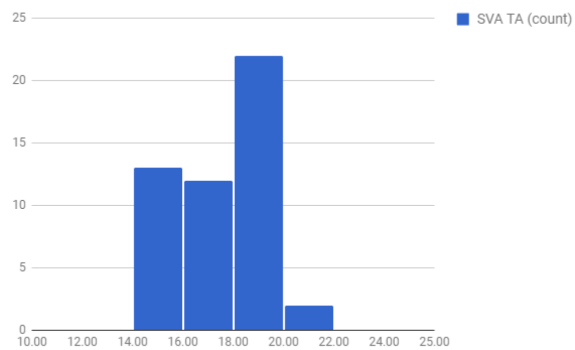


Figure 4.11. Histogram of SVA TA Scores

Individual responses were then rated using the HA rubric. During this rating, the writing was read holistically to gauge the overall effectiveness in argumentation. For the Newton's Law writing, the average HA score was 7.44, the median HA score was 7 the range of the HA scores was 5 and the standard deviation of the HA scores was 1.07. Figure 4.11 shows the histogram for the SVA writing HA scores.

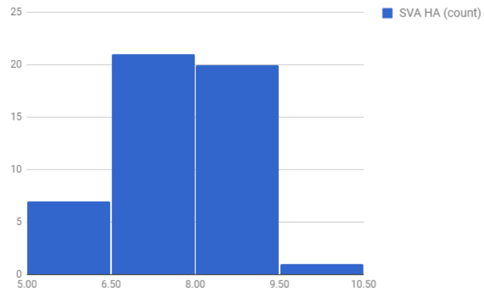


Figure 4.12. Histogram of SVA HA Scores

Simple Machines Writing

The Simple Machines Writing was the participants' fourth experience with the ACORN framework. The students had just completed several days of instruction on the six simple machines, mechanical advantage, work and power. For this writing assignment, the participants worked individually to answer the writing prompt. The researcher again modeled the use of the ACORN Framework as well as the think aloud protocol. Based on feedback from the students, a new ACORN graphic organizer was developed and used for this assignment (See Appendix). Participants addressed the problem in Figure 4.13 using the ACORN framework with the graphic organizer.

Observe the picture and quote below. Explain how this picture and quote by Archimedes is accurate based on how a lever works in terms of the position of the input (effort) force, output (resistance) force, load, fulcrum, and force-distance trade off. Remember to use ACORN when answering this question.

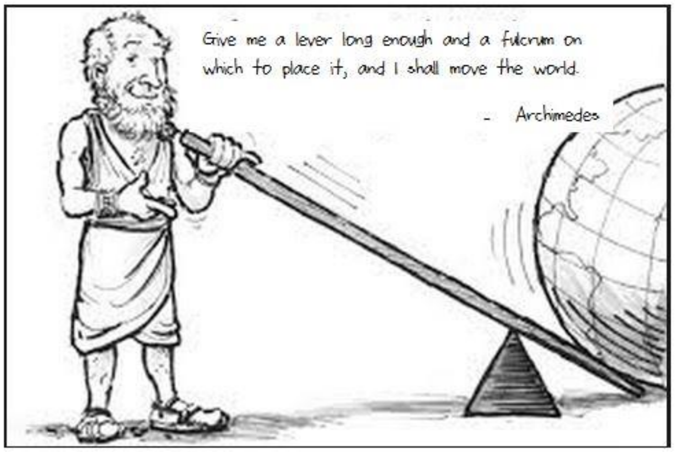


Figure 4.13 The Simple Machines Writing

Participants were not allowed to discuss the problem or the process of writing as a group. Each student was at his or her own computer and completed the writing using on the ACORN graphic organizer while using a microphone headset to capture the think aloud process. Each student had the freedom to frame their answer according to what they had learned.

Individual responses were first rated using the TA rubric. During this rating, each paper was marked showing instances of claims, evidence, and reasoning. Each instance was then compared the TA rubric and rated accordingly. For the simple machines writing, the average TA score was 14.80, the median TA score was 15, the range of the TA scores was 15 and the standard deviation of the TA scores was 3.77. Figure 4.14 shows the histogram for the simple machine TA writing scores.

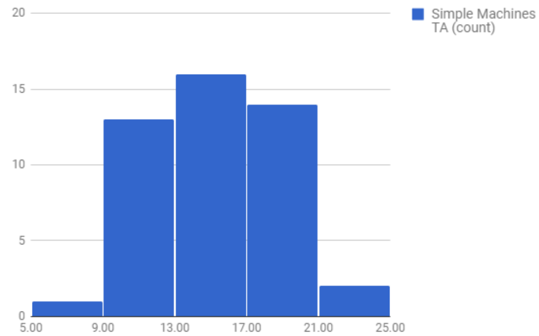


Figure 4.14. Histogram of Simple Machines TA Scores

Individual responses were then rated using the HA rubric. During this rating, the writing was read holistically to gauge the overall effectiveness in argumentation. For the Newton's Law writing, the average HA score was 4.63, the median HA score was 4.5 the range of the HA scores was 7 and the standard deviation of the HA scores was 2.1. Figure 4.15 shows the histogram for the simple machine HA writing scores.

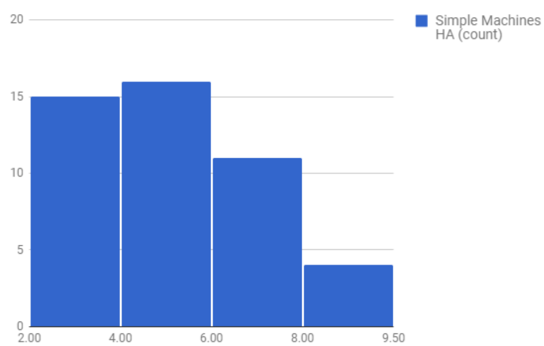


Figure 4.15. Histogram of Simple Machines HA Scores

Final Writing

The Final Writing was the participants' fifth experience with the ACORN framework. The students had just completed several days of application activities covering the entire unit. For this writing assignment, the participants worked individually to answer the writing prompt. The researcher again modeled the use of the ACORN Framework as well as the think aloud protocol. The students used the second graphic organizer to complete the writing assignment. Participants were given a list of eight topics to choose from in order to present an argument. The prompt for the task is shown in Figure 4.16

A trebuchet demonstrates several properties of physics. Choose one of the following physics properties that you feel is best represented in a trebuchet and explain how that property works.

- Gravity
- Newton's 1st Law
- Newton's 2nd Law
- Newton's 3rd Law
- Speed/Velocity/Acceleration
- Work/Power
- Simple Machine

In your writing, make sure you consider the following:

- A clear claim as to which property you are choosing
- A clear explanation of that property
- Evidence that supports your claim
- Reasoning that links your evidence to your claim.

Figure 4.16. The Final Writing Task

Participants were not allowed to discuss the problem or the process of writing as a group. Each student was at his or her own computer and completed the writing on the ACORN graphic organizer while using a microphone headset to capture the think aloud process. Each student had the freedom to frame their answer according to what they had learned.

Individual responses were first rated using the TA rubric. During this rating, each paper was marked showing instances of claims, evidence, and reasoning. Each instance was then compared the TA rubric and rated accordingly. For the final writing, the average TA score was 18.90, the median TA score was 19, the range of the TA scores was 17 and the standard deviation of the TA scores was 3.2. Figure 4.17 shows the histogram for the final writing TA scores.

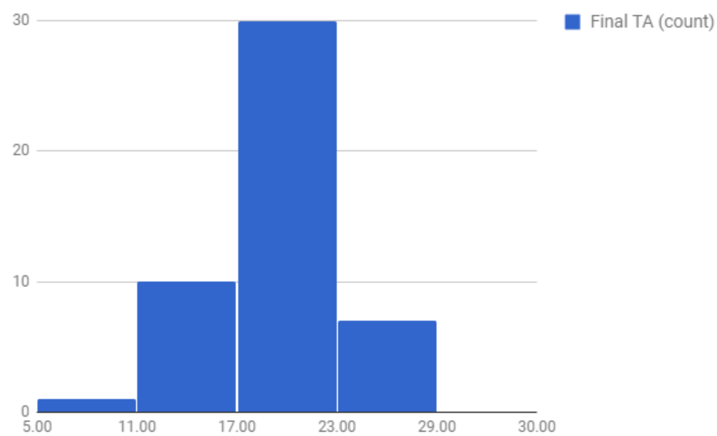


Figure 4.17. Histogram of Final TA Scores

Individual responses were then rated using the HA rubric. During this rating, the writing was read holistically to gauge the overall effectiveness in argumentation. For the Final writing, the average HA score was 6.71, the median HA score was 7 the range of the HA scores was 7 and the standard deviation of the HA scores was 1.6. Figure 4.18 shows the histogram for the final writing HA scores.

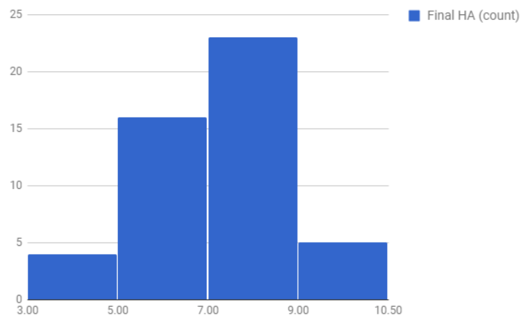


Figure 4.18. Histogram of Final HA Scores

Writing Samples Improvement

In order to see if students' responses improve overall from beginning to end, the average TA and the average HA scores were plotted according to each instance. Figure 4.19 shows the average TA score for each writing.

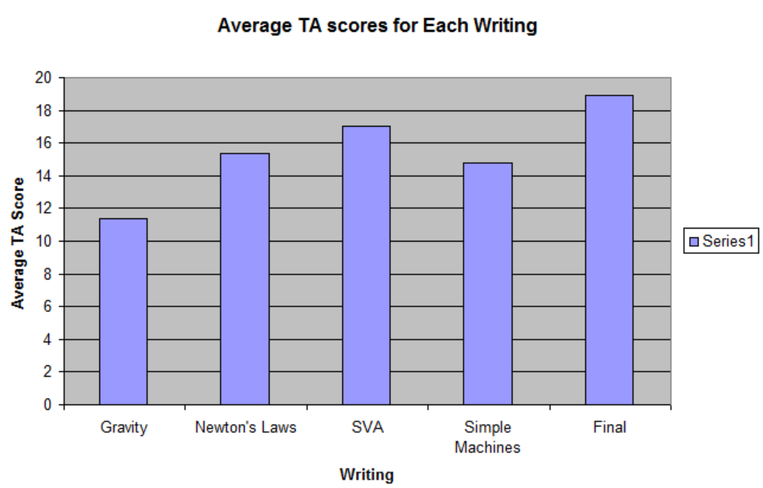


Figure 4.19. Average TA Score by Writing

With the exception of the simple machines writing, students made progress in their ability to produce written argumentation according the TA rubric. Discrepancies with the simple machine writing will be discussed in Chapter 5. Figure 4.20 shows the average HA score for each writing.

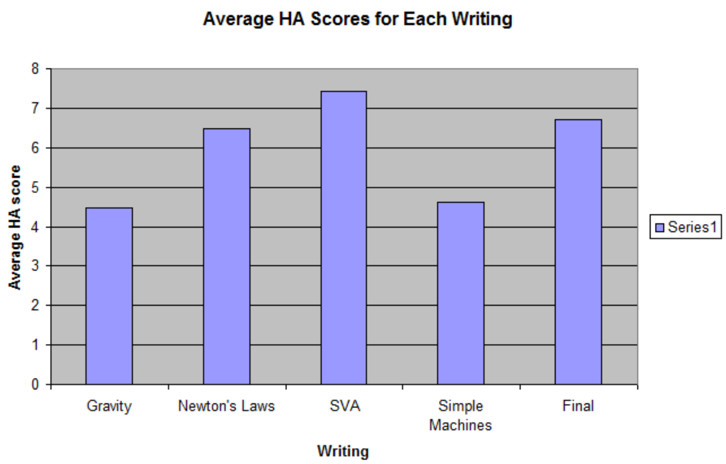


Figure 4.20. Average HA Score by Writing

With the exception of the SVA writing to the simple machines writing, and the SVA writing to the final writing, students made progress in their ability to produce written argumentation according the HA rubric. Discrepancies with the simple machine writing will be discussed in Chapter 5.

A paired t-test comparing the overall change from the gravity writing to the final writing was calculated. In the case of the improvement of the TA scores from the gravity writing to the final writing, there was significance ($p=1.7x10^{-20}$ and $\alpha=0.05$). In the case of the improvement of the HA scores from the gravity writing to the final writing, there was significance ($p=5.5x10^{-12}$ and $\alpha=0.05$). According to both rubrics students were able to improve the ability to produce a written argument.

Correlation between TA and HA scores

Using the two different rubrics to score the papers allows for two different views of argumentation (Choi, Notebaert, Diaz, & Hand 2010). In order to examine the relationship between the two rubrics, linear regressions were run between the TA scores and HA scores for each of the writing samples with the results in Table 4.1.

Table 4.1. Correlations Between TA and HA Scores for Each Writing

Writing	Correlation between TA and HA scores
Gravity	r= .64
Newton's Laws	r= .66
SVA	r= .39
Simple Machines	r= .81
Final	r= .76

The scores for each of the writings were moderately or strongly correlated to each other. This is shows that there is a relationship between being able to write structurally as measured by the TA rubric and holistically as measured by the HA rubric.

Student Writing Compared to the MAI

The primary writing of interest in the data set is the final writing because this is the writing where students have had the most experience and the most freedom in developing a written argument. Having the students do several iterations of writing is in line with the pragmatic theoretical framework of this research. In order to gain understanding of the relationship between the final writing and the students' metacognition at the time of the final writing, a linear regression was performed using both writing rubric scores and both elements of metacognition. Table 4.2 shows the results of the four linear regressions.

Table 4.2. Student Writing Scores Compared to Metacognition

	MAIK Post	MAIR Post
TA	r=.11	r=.05
HA	r=.23	r=.15

None of the final writing scores correlated strongly with metacognition as measured by the MAI.

Student Writing Compared to the FCI

In order to gain understanding of the relationship between the final writing and the students' subject matter learning at the time of the final writing, a linear regression was performed using both writing rubric scores and the FCI posttest score. Table 4.3 shows the results of the two linear regressions.

Table 4.3 Student Writing Scores Compared to Student Subject Matter Learning

	FCI Post
TA	r=-.02
HA	r=-.13

None of the final writing scores correlated strongly with subject matter learning as measured by the FCI.

MAI Compared to the FCI

The lack of correlation between writing and metacognition and writing and subject matter learning was unexpected. In order to investigate these results further, a linear regression comparing the FCI to each the MAIK post and the MAIR post was run. The results are found in Table 4.4.

Table 4.4. Metacognition Compared to Student Subject Matter Learning

	FCI Post
MAIK Post	r=-.03
MAIR Post	r=-.22

Student subject matter learning as measured by the FCI did not correlate strongly with metacognition as measured by the MAI.

Summary of Quantitative Results

Three main results were determined from the quantitative data. First, students demonstrated subject matter learning during the time of intervention according to the FCI. Second, students showed improvement in regulation of cognition during the time of the intervention. Third, students showed improvement in their ability to produce written argumentation during the time of the intervention.

Some of the unexpected results were the lack of correlation between and among the student writing scores, the MAI and the FCI. In order to investigate this further, a correlation matrix including all of the writings, the MAI post test scores, and the FCI post test scores was created as shown in Figure 4.21.

Correlations

Correlations

		TAFinal	HAFinal	MAIKPost	FCIPost	MAIRPost	TAGravity	HAGravity	TANewtons	HANewtons	TASVA	HASVA	TAMachines	HAMachines
TAFinal	Pearson Correlation	1	.762**	.107	-.021	.058	.355*	.374**	.386**	.331*	.097	.314*	.385**	.579**
	Sig. (2-tailed)		.000	.471	.888	.694	.013	.009	.007	.022	.510	.030	.008	.000
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
HAFinal	Pearson Correlation	.762**	1	.233	-.134	.147	.256	.349*	.409**	.379**	.170	.455**	.371*	.601**
	Sig. (2-tailed)	.000		.111	.364	.318	.079	.015	.004	.008	.247	.001	.011	.000
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
MAIKPost	Pearson Correlation	.107	.233	1	-.033	.776**	.168	.207	.202	.215	.037	-.076	.106	.026
	Sig. (2-tailed)	.471	.111		.822	.000	.253	.158	.168	.143	.802	.607	.481	.862
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
FCIPost	Pearson Correlation	-.021	-.134	-.033	1	-.218	.160	.092	.077	.045	.139	-.054	.011	.013
	Sig. (2-tailed)	.888	.364	.822		.136	.277	.533	.602	.762	.345	.714	.943	.934
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
MAIRPost	Pearson Correlation	.058	.147	.776**	-.218	1	.086	.072	-.004	.147	-.007	.051	.126	.099
	Sig. (2-tailed)	.694	.318	.000	.136		.563	.629	.981	.319	.961	.730	.403	.511
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
TAGravity	Pearson Correlation	.355*	.256	.168	.160	.086	1	.635**	.325*	.240	.064	.256	.350*	.382**
	Sig. (2-tailed)	.013	.079	.253	.277	.563		.000	.024	.101	.667	.079	.017	.009
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
HAGravity	Pearson Correlation	.374**	.349*	.207	.092	.072	.635**	1	.408**	.402**	.217	.228	.189	.285
	Sig. (2-tailed)	.009	.015	.158	.533	.629	.000		.004	.005	.138	.120	.210	.055
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
TANewtons	Pearson Correlation	.386**	.409**	.202	.077	-.004	.325*	.408**	1	.663**	.026	.224	.447**	.478**
	Sig. (2-tailed)	.007	.004	.168	.602	.981	.024	.004		.000	.859	.126	.002	.001
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
HANewtons	Pearson Correlation	.331*	.379**	.215	.045	.147	.240	.402**	.663**	1	.234	.282	.285	.488**
	Sig. (2-tailed)	.022	.008	.143	.762	.319	.101	.005	.000		.109	.052	.055	.001
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
TASVA	Pearson Correlation	.097	.170	.037	.139	-.007	.064	.217	.026	.234	1	.391**	.041	.209
	Sig. (2-tailed)	.510	.247	.802	.345	.961	.667	.138	.859	.109		.006	.785	.164
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
HASVA	Pearson Correlation	.314*	.455**	-.076	-.054	.051	.256	.228	.224	.282	.391**	1	.466**	.566**
	Sig. (2-tailed)	.030	.001	.607	.714	.730	.079	.120	.126	.052	.006		.001	.000
	N	48	48	48	48	48	48	48	48	48	48	48	46	46
TAMachines	Pearson Correlation	.385**	.371*	.106	.011	.126	.350*	.189	.447**	.285	.041	.466**	1	.810**
	Sig. (2-tailed)	.008	.011	.481	.943	.403	.017	.210	.002	.055	.785	.001		.000
	N	46	46	46	46	46	46	46	46	46	46	46	46	46
HAMachines	Pearson Correlation	.579**	.601**	.026	.013	.099	.382**	.285	.478**	.488**	.209	.566**	.810**	1
	Sig. (2-tailed)	.000	.000	.862	.934	.511	.009	.055	.001	.001	.164	.000	.000	
	N	46	46	46	46	46	46	46	46	46	46	46	46	46

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4.21. Correlation Matrix of Quantitative Results

As shown in the matrix, each writings' TA and HA scores had either a moderate or strong correlation noting the consistency in the scoring of the writing. Also, MAIK post and MAIR post were strongly correlated (r=0.776) suggesting that the expected link between knowledge of cognition and regulation of cognition is present. No other correlations existed which will be discussed in Chapter 5.

Qualitative Data Sets

Two different qualitative data sets were examined as follows: student think alouds and student interviews. Each provides an alternative view of the intervention and will be used to give insights to the quantitative data (Cresswell, 2014; Greene, 2007).

Think alouds

The think alouds were used to capture elements of student metacognition. Overall, the students participated in five different think aloud sessions, one corresponding to writing. Each time, the think aloud protocol was modeled at the beginning of class. The purpose for having students participate in multiple think aloud sessions was primarily to have the students become familiar with using the microphones and computer to record themselves as well as gain a comfort level with talking aloud what they are feeling.

While all of the sound files associated with all of the writings were saved, only specific files were selected to be analyzed. Only sound files from the final writing were selected, under the assumption that by that time students were the most comfortable writing using the ACORN framework and recording themselves. Also, using only the final sound files coincides with my decision to use the final writings for analysis.

Nine students' sound files were selected for the qualitative analysis. To select the students' files, the following procedure was followed. First students' post scores for the MAI were plotted versus the TA scores for the final writing as shown in Figure 4.22

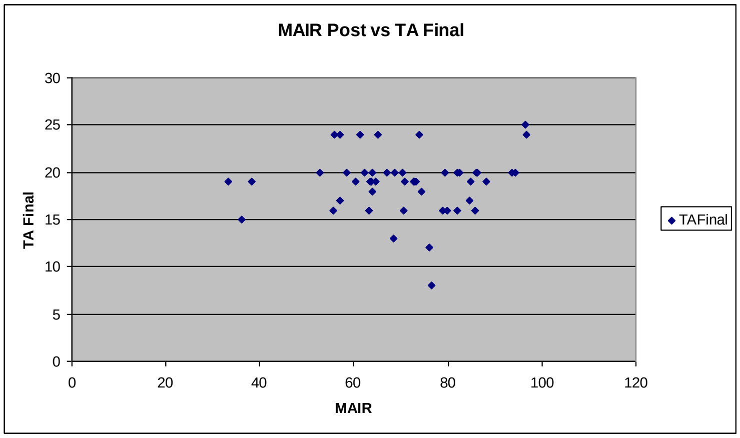


Figure 4.22. Student Regulation of Cognition Post Test Scores vs. Final Writing TA Scores

Next, the scatter plot was divided into four quadrants, the axes defined by the average score. Then, students in the extremes of each quadrant were selected. This yielded eight students. The process was repeated replacing the TA scores for the final writing with the HA scores for the final writing shown in Figure 4.23.

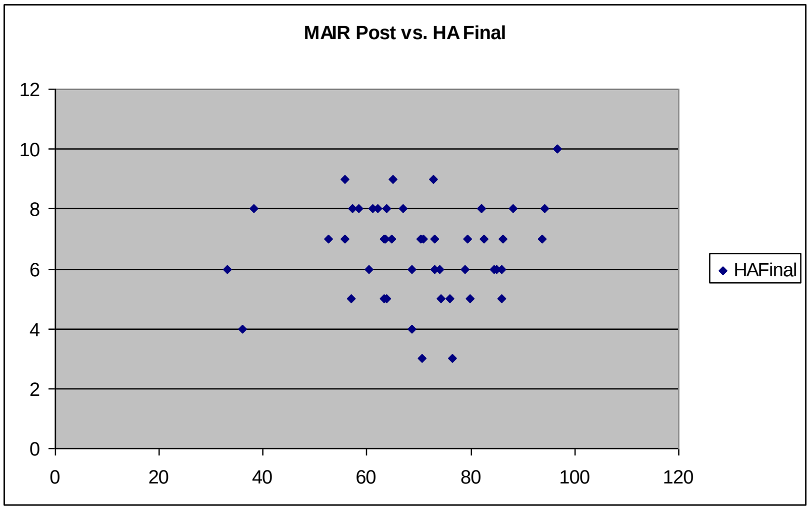


Figure 4.23. Student Regulation of Cognition Post Test Scores vs. Final Writing HA Scores

This yielded eight students as well, with seven of the students in common from the TA selection. I chose to include all students' files that resulted from either search in order to include as much variety as possible. Each sound file was then transcribed using an audio player and Microsoft word.

By selecting students in each quadrant, I can look the results by performance level defined by students who fall in the following quadrants: high metacognitive, high scoring (HSHM) students; high scoring, low metacognitive (HSLM) students; low scoring high metacognitive students (LSHM); and low scoring, low metacognitive (LSLM) students. HSHM, LSLM students are the expected outcome while HSLM and LSHM students are the exception.

Once the sound files were selected and transcribed, analysis began by identifying instances of both categories of knowledge of cognition and regulation of cognition. In order to help define each of the categories, I cross-referenced the verbiage of the questions on the MAI with the each of the subcategories listed in Table 4.6. This was easy to do since the MAI is also subdivided into the same subcategories. A definitional list was developed a priori in order to guide the analysis process, but the definitional list was not used to exclude words, phrases, and utterances from the data set. Analysis proceeded according to the tenets of the CCM (Glaser, 1965) in which the transcripts were analyzed by comparing possible instances of metacognition with the definitional list and the data in each transcript. The definitional list can be found in Table 4.5.

Table 4.5 Definitional List Used to Guide the Identification of Instances of Metacognition

Type	Sub Category	Definitions
Knowledge of Cognition		
	Declarative (DK)	• Understands intellectual strengths and weakness • I know what is important to learn • I am good at organizing/remembering • I know what I am expected to learn • Control over learning • Can judge undertstanding • Interest tied to learning
	Procedural (PK)	• Used what works in past • Purpose for strategies • Awareness of strategies • Automatic use of strategies •
	Conditional (ck)	• Learn best when have prior knowledge • Different strategies according to situations • Can motivate myself when needed • Use strengths to compensate for weaknesses

		• Know when strategies will work •
Regulation of Cognition		
	Planning (p)	• Pacing for enough time • Preplanning • Set goals before task • Ask questions before beginning • Read instructions before beginning • Organize time to accomplish goals
	Monitoring (m)	• Meeing goals • Considering alternatives before answering • Checking for all alternatives after answering • Review to help understand relationships • Analyzing usefulness in strategies • Asks myself how I am doing
	Evaluation (e)	• How did I do after complete • Asks if easier way • Summarizes • Asks if reached goals when done • Ask if considered all options when done • Asks if learned as much as could have •

After the sound files were transcribed, the qualitative data was transformed into a quantitative format (Greene, 2007, Creswell, 2014). Each transcription was analyzed for instances of metacognition. Specifically, words, phrases, and other utterances, were identified and categorized according the instances of knowledge of cognition and regulation of cognition. Each instance was coded with one of the following: PK, DK, CK, P, M, E. Table 4.6 shows examples from the transcripts as well as the total number of instances in all nine transcripts. In addition to instances of metacognition, I was also interested in whether or not students engaged in the ACORN framework. Direct references to ACORN were also noted.

Table 4.6. Examples and Number Instances of Metacognition

Type	Sub Category	Examples	Number of instances
Knowledge of Cognition	Declarative (DK)	"we got to answer a question about making a claim"	17
	Procedural (PK)	"this is like when we did that, wrote about the three laws and the trebuchet"	12
	Conditional (CK)	"because we actually spent so many days working on it"	1
Regulation of Cognition	Planning (P)	"Okay, now for another evidence"	111
	Monitoring (M)	"I think I'm going to write about the first law"	204
	Evaluation (E)	"you know what I'm just not going to put very much of a conclusion."	29
References to ACORN		"Cite evidence to help support your claim"	41

In summary these nine students produced a total of 374 instances of metacognition. The overwhelming majority of instances fell into the Regulation of Cognition category with a total of 344 for all nine students. Table 4.7 provides a different view of the data by organizing the information by student.

Table 4.7. Student Pseudonyms, Instances of Metacognition, and Performance Level

Pseudonym	Instances of Metacognition	Performance Level
Nancy	56	HSHM
Laura	23	HSHM
Enid	82	HSLM
Carlos	43	HSLM
Andrew	40	LSHM
Erica	73	LSHM
Natasha	29	LSLM
Keith	2	LSLM
Jerry	67	LSLM

Organizing the information in this manner helps us to see the difference in instances of metacognition as it relates to performance level. Although statistical analysis with N=9 would not provide meaningful results, we can note the trend that individual instances of metacognition during think alouds does not necessarily match results from the MAI. Jerry had a low MAI score but had higher instances of metacognition in the think aloud. Likewise, Laure had a high MAI scores, but had lower instances of metacognition in the think aloud.

While these transcriptions provide rich data for understanding how middle school student think about writing, the analysis of the think alouds was limited to instances of metacognition in line with research question number two. Anecdotal information in support of research question number three will be introduced as warranted in Chapter 5.

As a different view of the above data, Table 4.8 shows the average number of instances for each performance level. The students whose transcripts were analyzed were grouped into HSHM, LSHM, HSLM, and LSLM. The average score for knowledge of cognition (KOC) and regulation of cognition (ROC) was then calculated for the students in each grouping.

Table 4.8. Instances of Metacognition by Performance Level

Metacognition vs. Scoring	LM	HM
HS	KOC=1.5 ROC=16.67	KOC=.67 ROC=15.8
LS	KOC=1.1 ROC=8.9	KOC=1.2 ROC=15.8

With the small data set (n=9) traditional statistical analysis can not be performed, but some trends can be noted. First, HSHM students showed the lowest number of instances of knowledge of cognition. Second, only the LSLM students demonstrated a low number of instances of regulation of cognition when compared to the other groups. Given the results of the quantitative data sets, these results provide insight which will be discussed in Chapter 5.

Student Interviews

Ten students were selected to participate in the interview process. These students were chosen based on their perceived performance on the writing tasks. In other words, these students were chosen to represent a wide range of student performance based on my perception of how they completed the tasks in class, the quality of their work, and their behavior when engaging with the ACORN rubric. One of the advantages of being the teacher of the research subjects was having a solid understanding of the students' ability in writing, argumentation, and student learning. Being able to monitor how the students engaged with the rubric, how they wrote on a variety on assignments, and how well they were learning the information guided me in my choices. Also, being the teacher, I have established a rapport with the students which was useful during the interview process.

I attempted to include students that demonstrated different performance levels. I was able to determine by monitoring and assessing the students daily in the context of the classroom setting. It was my assessment of these students as a teacher, not a researcher, that lead to their selection as interview participants. These students were purposely selected not based on scores on writing, the FCI, or the MAI. Also, students were selected before any of the quantitative or qualitative data analysis was done. This was done as a possible additional source of triangulation. Students selected based on perception gives another view to the data set and allows me to compare the data in the intervention to my perceptions of the intervention. After the students were selected and interviewed, their position on Figure 4.19 was located and each student was identified according to their performance level. Student pseudonyms and their performance level can be found in Table 4.9. Table 4.9 also shows whether or not their think aloud was transcribed and analyzed.

Table 4.9. Student Pseudonyms for Interviews and Performance Level

Student Pseudonym	Performance Level	Think Aloud Analyzed
Larry	HSHM	No
Braden	HSLM	No
Jerry	LSLM	Yes
Joy	LSLM	No
Nancy	HSHM	Yes
Elizabeth	LSHM	No
Trace	HSLM	No
Meg	HSHM	No
Linden	LSLM	No
Amy	HSHM	No

Of the ten students selected, all four performance levels were represented, albeit the LSHM quadrant was only represented by one student. Since these students were preselected and ultimately ended up representing all four performance levels supports the idea that the perceived performance of the students matches the actual performance of the students.

All ten participants agreed to participate and agreed to have the interview recorded. The semi-structured interview was conducted at school during their science class time. I transcribed the interviews using an audio player and Microsoft word.

The interviews were analyzed based on the tenets of the CCM (Glaser, 1965). Each interview was transcribed and then responses were organized by interview question. After this, the responses were compared to each other through the process of several readings. As the comparisons progressed, five themes were identified from the interviews: ACORN as an organizer, ACORN promotes metacognition, ACORN helps in subject matter learning, Confusion about the Steps of ACORN, ACORN helps differentiate between evidence and reasoning. Each of the five themes is discussed below.

ACORN as an organizer

One of the main ideas to come from the student interviews was that the students saw the ACORN process as an organization tool as opposed to a learning tool. All of the students either directly indicated that the ACORN framework was more for organization than learning or made reference to the ACORN framework as an organization tool. Furthermore, there was little or no hesitation when posed with this question. Larry said, "I though it was pretty good because it gave a good way to organize your thoughts and ideas and stuff." Joy echoed this idea when talking about ACORN as an organizational tool and referred to it as "a checklist in my head and get straight to the point."

Students also indicated that ACORN was a better organizational tool than using the RACE format. Students appreciated the steps defined by the acronym as a way to put together their thoughts. Braden states that ACORN "allows for an extra step organize over the normal RACE". Throughout the interviews, students continued to refer to the various steps and allowing them to put together their thoughts, although there was not one specific step that students could agree upon as the most useful for this purpose. Other than the Answer step, Cite, Organize, Respond, and Note were all discussed as ways student thought about organizing the evidence.

ACORN promotes metacognition

When asked whether or not the ACORN framework made them think about what they were writing, students said yes. Also, students referred to going back and thinking about writing as a part of the process. These types of thoughts align closely with regulation of cognition. For example, planning was evident when Trace said, “I first had to think about why, then I had to collect my evidence.” Monitoring was evident when Larry states, “Basically I kind of thought what I was going to write next if it was going to make sense while I was writing. If it was going to be good enough.” Evaluation was the most common reference in the interviews, particularly when talking about the Note. Amy describes the Note as allowing her to think about her writing “and you could go back and fix what you wrote”. Overall all of the students indicated at some point that the steps of ACORN caused them to think about their writing.

Also important to note is that very few students displayed knowledge of cognition in their interviews. Few comments were made that suggest knowledge of cognition such as when Nancy said, “I look at my past writing and see what I did” indicating procedural knowledge of cognition, but overall student comments showed regulation of cognition.

ACORN helps in subject matter learning

First, students indicated that the writings using the ACORN framework related to the activities done in class. Students expressed no confusion on this point which is important because the activities in class are the main point of instruction. The activities were designed to instruct students on content. That the writings were related to the activities establishes the first link to what is done in class and the subject matter learning. No students hesitated or expressed any concerns that the writings were not a part of the activities.

Second, students indicated that writing using ACORN helped them to learn the content. Meg states, “I think it helped me sometimes because I was confused about what exactly we were talking about. I could sort through under like my evidence and how everything pieced together”. Most of the students indicated similar notions about how the writing process helped them to make sense of the content. According to the students interviewed, putting together the information in an organized format is what aided in their subject matter learning. For example, Larry states, “like when you get evidence from the book, you might not understand it, but if you can ask questions like what is an example, you can write it down on paper, it helps you learn it.” Students indicated that the ACORN framework helped with that sense making process.

Confusion about the Steps of ACORN

Almost all of the students indicate some initial confusion about the steps in the ACORN framework. None of the students indicated that the Answer or Cite steps were an issue which is interesting to note because these two steps are also found in the RACE framework. Confusion about Organize, Respond, and Note, seemed to be equally spread among the students. For example, Jerry expressed his concern that the Organzie and Respond step were out of order. Nancy did not see the purpose of the Note because “I do not have to write out what I already did.” Most of the students also indicated an initial confusion about the steps, but also indicated that after working with the framework that it got better. Amy states, “It was kind of confusing at first, it was different from RACE like format, but after we wrote like four times or so it got easier to use and I feel like it actually did help more.” A part of the confusion was related to the students being familiar with RACE and then having to learn a new framework.

ACORN helps differentiate between evidence and reasoning.

One of the points that all of the students made was how the ACORN framework helped them to differentiate between evidence and reasoning. This is important because the issue of distinguishing between different aspects of argumentation is an ongoing problem with students and argumentation (Sampson & Clark, 2008, Sandoval & Millwood, 2007). Braden highlights this idea when he says, "I think so, RACE kind of lumps them together, but ACORN really separated them. I could see what was evidence and what was reasoning." Most students echoed this idea except for one student Jerry who said "no I don't think it did. Because with CER it was more structured". In this case Jerry stated his preference for the CER framework since it was explicit in differentiating evidence from reasoning.

Even when not answering a direct question about evidence and reasoning, students would indicate that ACORN helped them. When talking about what he liked best about ACORN, Linden replied, "I would have to see which order I would have to put evidence in and my reasoning in" Other students replied similarly indicating that the ACORN process helped them to organize and distinguish between evidence and reasoning.

Summary of Qualitative Data

In an Explanatory Sequential Design, quantitative data are collected and analyzed and then qualitative data is used help explain the results of the quantitative analysis (Creswell, 2014). The think aloud data set helps to explain the results in the quantitative set showing the significance in the increase of regulation of cognition. In the student think alouds, students demonstrated more instances of regulation of cognition then knowledge of cognition.

As for the student interviews, five themes were identified. The two themes of ACORN promotes metacognition and ACORN helps in subject matter learning relate directly to the quantitative data sets showing the increase in regulation of cognition as well as the increase in subject matter learning as measured by the FCI. The theme of ACORN helps differentiate between evidence and reasoning is important because it addresses a longstanding issue in the field of argumentation. The two themes of Confusion about the Steps of ACORN and ACORN as an organizer, provide possible foci for future research.

Results by Research Question

Research Question 1

Using the ACORN framework, a student can produce a high scoring, structured argument as measured by the TA Rubric or the HA Rubric. This is supported quantitatively. Since enough samples were taken for each writing instance (N=48), I would expect to see a normal distribution of scores for each of the writings and for each rubric. All five writings, each scored by both rubrics showed a normal distribution in all ten cases matching my expectations.

The purpose of using the TA and HA rubrics is to allow for two different viewpoints of written argumentation. For the ACORN framework to be considered a form of argumentation, the basic elements of claims, evidence, and reasoning must be present (Bricker & Bell, 2008). The TA rubric was designed to look at the different components of argumentation, mainly the question, the claim, the evidence, the reasoning and the relationship among those components (Choi, Notebaert, Diaz, & Hand 2010). The benefit of using the TA rubric is that it considers not only each element separately, but also the relationship between the elements. This rubric was used specifically because it addressed the structural approach to argumentation (TAP) and the widely accepted idea that argumentation in science education is evaluated in this manner (Bricker & Bell, 2008). Papers scored according to this rubric demonstrate that a student can produce the elements and structure of argumentation.

In Figure 4.24, Braden demonstrates each of the elements of claims, evidence, and reasoning in his final writing response. His score was average at 19 with scoring a 4 in each area of the rubric(see and example of the rubric in the Appendix) with the exception of quality of evidence which was a 3.

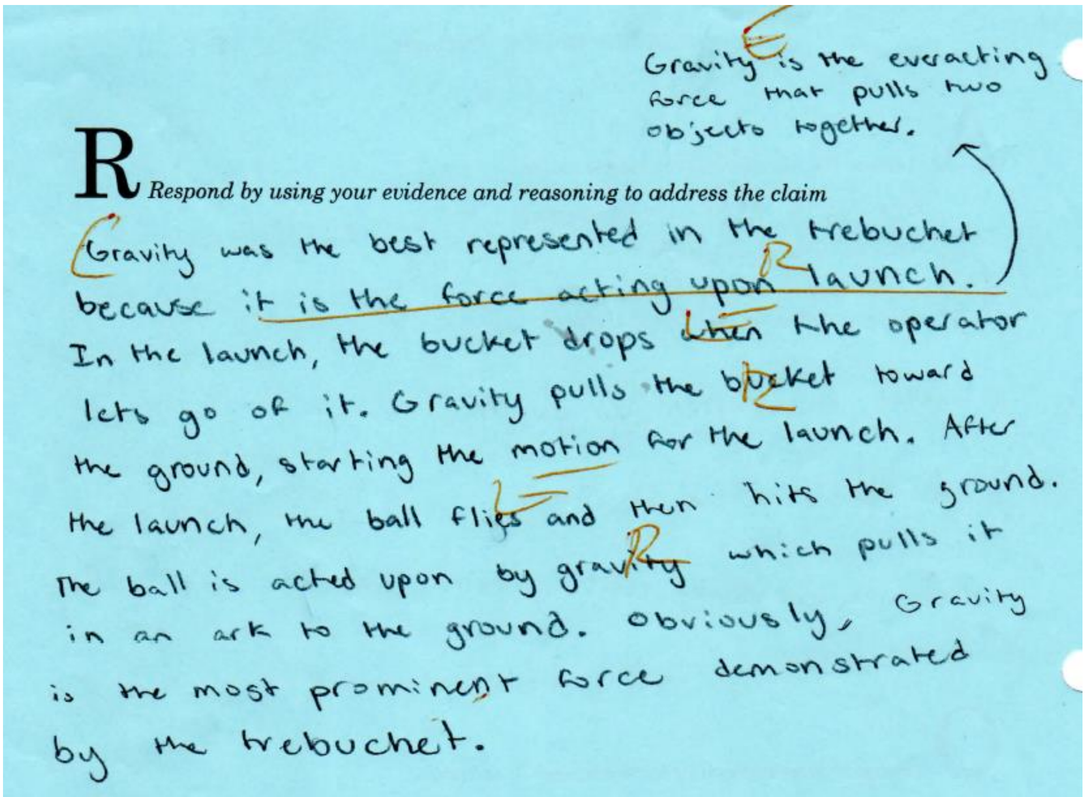


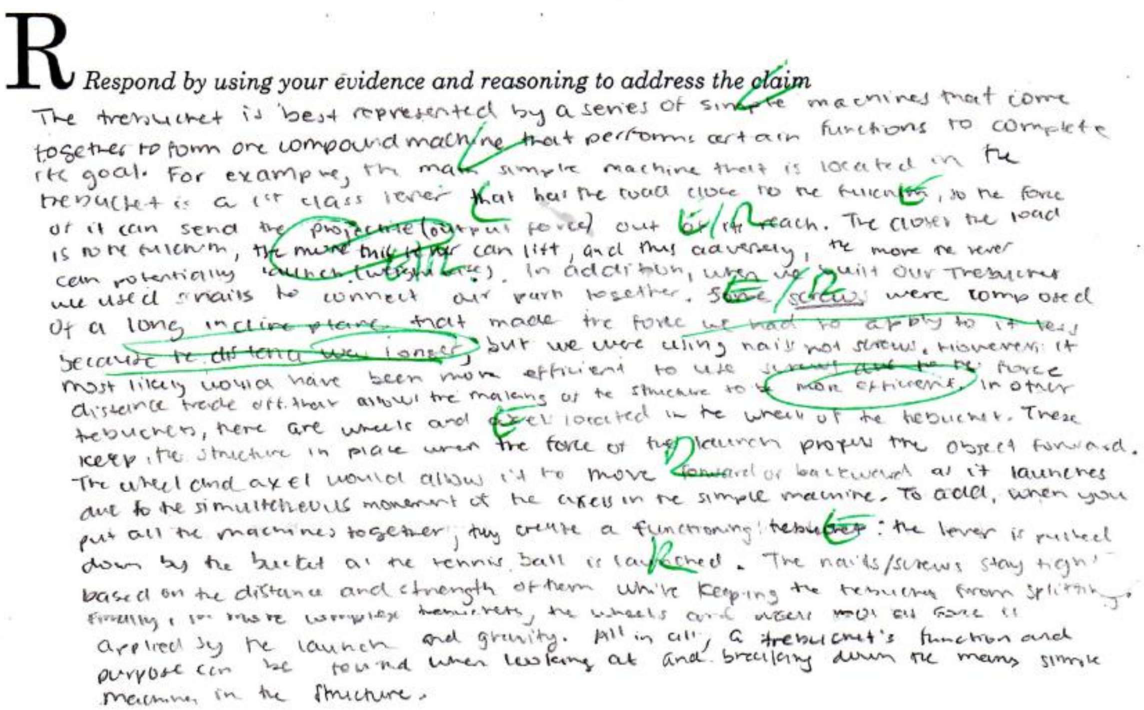
Figure 4.24. Example Showing Elements of Claims, Evidence, and Reasoning

Included in the written response is the rater's designation of claim (C), evidence (E), and reasoning (R).

To successfully employ the rubric to get a wide variety of scores, each element of argumentation would have to be present. Since a normal distribution on the TA rubric for each of the writings was present, the participants were able to produce a variety of writings demonstrating different levels of written argumentation. There would be no high-end scores if

ACORN did not include all aspects of argumentation. The presence of the high-end scores in the distribute shows that at least some students can use ACORN to supply the necessary components to create a solid scientific argument.

In Figure 4.25, Laura presents a high scoring argument on her final writing with a score of 24. She received a 5 in each category of the rubric with the exception of reasoning.


R *Respond by using your evidence and reasoning to address the claim*

The trebuchet is best represented by a series of simple machines that come together to form one compound machine that performs certain functions to complete its goal. For example, the main simple machine that is located in the trebuchet is a 1st class lever that has the load close to the fulcrum, so the force of it can send the projectile (output force) out further reach. The closer the load is to the fulcrum, the more this lever can lift, and thus adversely, the more the lever can potentially launch (output force). In addition, when we built our trebuchet we used nails to connect our parts together. Some/screws were composed of a long inclined plane that made the force we had to apply to it less because the distance was longer, but we were using nails not screws. However, it most likely would have been more efficient to use screws and for the force distance trade off that allows the making of the structure to be more efficient. In other trebuchets, there are wheels and axel located in the wheels of the trebuchet. These keep the structure in place when the force of the launch propels the object forward. The wheel and axel would allow it to move forward or backward as it launches due to the simultaneous movement of the axel in the simple machine. To add, when you put all the machines together, they create a functioning trebuchet: the lever is pulled down by the bucket as the tennis ball is launched. The nails/screws stay tight based on the distance and strength of them while keeping the trebuchet from splitting. Finally, in more complex trebuchets, the wheels and axels move as force is applied by the launch and gravity. All in all, a trebuchet's function and purpose can be found when looking at and breaking down the many simple machines in the structure.

Figure 4.25. Example Showing Elements of Claims, Evidence, and Reasoning

Again, the rater's marks are included. Laura includes very strong reasoning, but according to the rubric, she did not reflect on new things to think about and she did not include a real life example outside of the trebuchet. These two elements have been found in other students' writings indicating that a perfect scoring paper can be produced.

The purpose of the HA rubric is to examine a different viewpoint of argumentation based on the "extent to which the whole argument was considered to be coherent and powerful." (Choi, Notebaert, Diaz, & Hand, 2010, p. 155). In this scoring, individual components of argumentation were not considered, but the overall effectiveness of the argument. Holistically speaking, a convincing scientific argument cannot be present if the argument is not coherent and powerful (Choi, Notebaert, Diaz, & Hand, 2010). The following student writing is the complete response of the lowest scoring final writing according to the HA rubric.

> A simple machine is represented in the trebuchet. A lever is one of the simple machines on the trebuchet. The throwing arm and the metal bar would be the parts on the lever. The metal bar would be the fulcrum.

This argument is neither powerful nor coherent given the age and grade level of the student although it does include a claim, evidence, and reasoning. In other words, even though structurally speaking, an argument is present, holistically speaking, this is not a good argument. Overall, students were able to produce a variety of scores including high-end scores to show that ACORN can be used to promote high quality written argumentation according to the HA rubric.

One of the benefits of scoring holistically is that it allows for a student to show argumentation skills without being bound the components of TAP. Students are allowed a broader scope in presenting an argument. While students who are able to produce a solid argument with a high scoring TA score typically also do well on the HA rubric it is possible to score well on the HA rubric, but not the TA rubric. Consider the following response from a student who score an 8 on the HA rubric, but a 16 on the TA rubric. The insertion of claim, evidence, and reasoning from the rater were included in parentheses.

> I believe that gravity is best represented in a trebuchet (claim). When building a trebuchet, we observed that the counterweight drops to the ground when we let go of the longarm (evidence). This is because gravity is a universal force that acts on all objects at all times, but it is very weak (reasoning). This is why we were able

> to act against it and bring the weight to the ground (reasoning). However, when we let it go (in an attempt to launch our projectile,) gravity acted upon it and pulled it down (evidence). Something else I observed when creating our trebuchet is that on average, the more sand we had in the bucket, the farther are projectile went (evidence). This is because there is more gravity between more massive objects (reasoning).

Given the level of the student, this is holistically a powerful and coherent argument as to why gravity is best related to a trebuchet. The response includes grade level appropriate understandings of gravity, how gravity was related to the actions of the trebuchet and was written in a logical order. Structurally, the quality of the claim, the quality of the evidence, the relationship between the claims and the evidence, and the reasoning were all deemed to be moderate. Each instance of evidence and reasoning missed the powerful level by not extending the ideas. For example, in the last set of evidence and reasoning, the student is entirely correct in what he said for evidence and reasoning, but missed the powerful level by not extending the ideas. In the last set of evidence and reasoning, the student is entirely correct in what he says, but misses the connection to Newton's second law that would have pushed the argument into the strong powerful level.

Overall, there is a moderate or strong correlation between the TA and HA rubrics for each writing. This shows that while there is a correlation between being able to write with structure and write holistically, but being able to write one way does not preclude being able to write the other way as demonstrated in the example above. On the other end of the argument, students who write structurally or holistically should be able also write holistically or structurally respectively. This is demonstrated by the fact there is a moderate or strong correlation. This is why using both rubrics to score the papers allows for a wider look at how students present argumentation thus lending strength to the argument to the use of ACORN is able to produce high quality argumentation.

Research Question 2

Students primarily engage in metacognition through use of regulation of cognition. This is supported both qualitatively and quantitatively. Qualitatively, nine students engaged in 344 instances of regulation of cognition compared to 30 instances of knowledge of cognition. This is also supported quantitatively by the significant improvement in the MAI regulation of cognition scores ($p=0.006$ and $\alpha=0.05$) while there was no significant improvement in the MAI knowledge of cognition scores. In other words, qualitative and quantitative data shows convergence (Greene, 2007) on this matter.

Regulation of cognition is the process of controlling learning through planning, monitoring, and evaluation. Students who display these traits display metacognition through a specific set of activities. Planning involves strategies pacing for enough time, preplanning, setting goals before a task, asking questions before beginning, reading instructions before beginning, and organizing time to accomplish goals. Typically, students performed these activities at the beginning of their writing when considering how they would answer the question. For example, planning was captured when Enid in her think aloud made the following statements, “Okay, so, what are the choices? (P) Uhh, I think I will do either (P) second law, or simple machines”. In the first instance, the student is asking questions before beginning and in the second instance they are setting goals.

Monitoring involves the set of strategies used to assess progress as work is progressing which includes meeting goals, considering alternatives before answering, checking for all alternatives after answering, reviewing to help understand relationships, analyzing usefulness in strategies, asking myself how I am doing, For example, Carlos states, “ok I don’t think I (M) should say anything about it being in a parabola, because that’s really not related to the law”. In this case, Carlos is considering alternatives before answering. In another example from Nancy, monitoring alternatives after answering is seen, "I'm forgetting, (M) why am I talking about a hammer." This type of metacognition was the by far the most prevalent in the student think alouds.

Evaluation, the third type of regulation of cognition involves activities such as asking how did I after completing, asking if there was an easier way, summarizing, asking if reached goals when done, asking if considered all options when done, and asking if learned as much as could have. For example, Jerry states, "Trebuchet, all right we have all the things we're talking about (E)." This shows if Jerry was considering all the options when he was done. Typically, evaluation occurs at the end of writing and was the least displayed form of regulation of cognition among the nine students' think alouds.

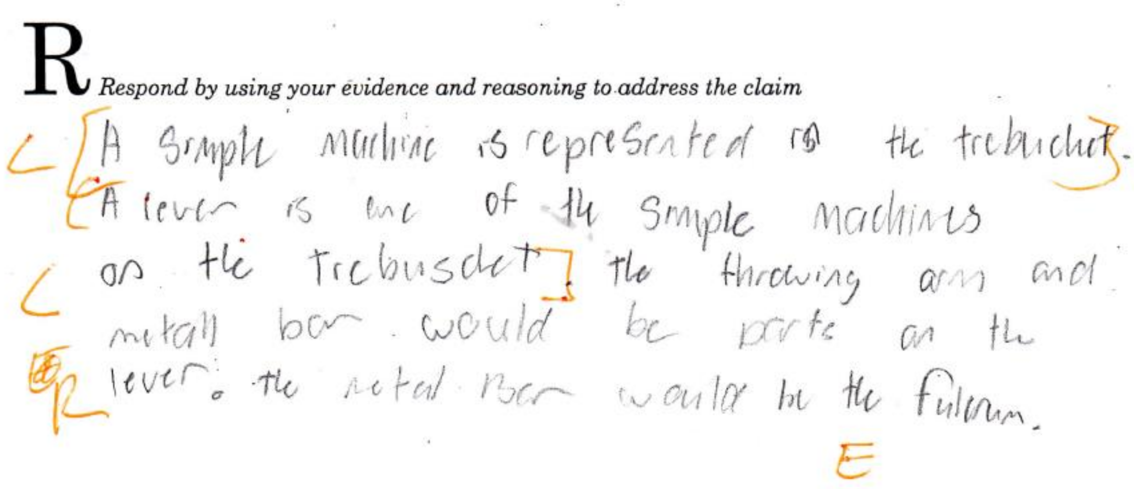
Qualitative data from the student interviews also show that students engaged in regulation of cognition. The primary examples were given under the qualitative analysis section for the student interviews. The students interviewed gave examples that related to the regulation of cognition. In other words, the students interviewed most often stated that the ACORN rubric helped them to plan, monitor and evaluate their writing. Given this was a goal of the ACORN framework, having qualitative data in the form of student interviews and think alouds as well as quantitative data from the MAI converge is an important result.

Research Question 3

Subject matter learning was examined both quantitatively by the FCI and qualitatively from the writing samples, think alouds and student interviews. Neither scores from the TA rubric nor scores from the HA rubric correlate to the FCI post test scores (r= -0.02 for TA, r= -0.13 for HA). There is no quantitative evidence to support that using the ACORN framework promotes student subject matter learning. There is qualitative evidence from the student interviews, the writing samples and the think alouds that indicate that the ACORN framework did aid in student subject matter learning though. In other words, qualitative and quantitative data diverges on this matter (Greene, 2007).

Four examples of students' final writings were analyzed. I have selected the following examples because they represent a low TA score, two average TA scores, and a high TA score. In each case, a different level of student learning is evident which does correspond to the TA score.

In the first example scored an 8 on the TA rubric. Figure 4.26 shows the entirety of the student writing.


R Respond by using your evidence and reasoning to address the claim

A simple machine is represented in the trebuchet. A lever is one of the simple machines on the trebuschet. the throwing arm and metall bar would be parts on the lever. the metal bar would be the fulcrum.

Figure 4.26 A Low Scoring Response

In this example, little subject matter learning is demonstrated. The student does correctly identify a lever as a part of trebuchet as well as identify the fulcrum, but nothing else. In other words a low scoring TA does correspond to a lack of subject matter learning.

The second example scored a 19 on the TA rubric which is an average score. In Carlos's final writing, he argues how gravity affects the function of a trebuchet. In one claim-evidence-reasoning scenario, he states "Gravity causes the ball to come down (C) because it applies a force to the ball (R) causing it causing it to change the direction that it was put into by the throwing arm (E)." Carlos is specifically referring to the concept of trajectory which was a topic explicitly taught in class. That Carlos can apply trajectory to the workings of his trebuchet indicates student learning. Furthermore, Carlos has learned more than he writes on the paper. The following is from Carlos's think aloud that corresponds to his written statement. In this case, the letters in the parenthesizes relate to regulation of cognition.

> I'll explain that one last (P). The ball comes down, the ball down gravity a force on it and that changes the direction of the ball when the initial force that sent the ball up becomes weaker than the force of gravity. OK lets read this again (M), The ball comes down, the ball down gravity a force on it and that changes the direction of the ball when the initial force that sent the ball up becomes weaker than the force of gravity. That makes sense (E)

Carlos's thinking is more detailed than his writing and indicates a better understand of the subject matter than his final writing. Also Carlos metacognitively engaged in regulation of cognition through planning (P), monitoring (M), and evaluation (E). This is one example that shows how the use of argumentation fosters metacognition in student learning.

The third example also scored a 19 on the TA rubric which is an average score. In an example of Newton's Third Law, Elizabeth uses evidence and reasoning. She states, "when the tennis ball launched and hit the ground (E), and the ground pushed back on it (R) causing the ball to bounce." The concept of the ground pushing back is not a natural one for a student since no one sees the ground pushing back. That Elizabeth can articulate that concept shows student subject matter learning. In Elizabeth's interview, she said that using the ACORN process caused you to have to cite information from what had been done in class. She also indicated that the ACORN process helped her to think about what she had learned indicating metacognition.

Laura also shows subject matter learning in her final writing. This writing earned a score of 24 on the TA rubric. In this claim-evidence-reasoning scenario, Laura discusses how a lever is found in a trebuchet. "The main simple machine that is located in the trebuchet is a 1st class lever (C) that has the load close to the fulcrum (E), so the force of it can send the projectile (output force) out of its reach (R)." The workings of a first class lever were explicitly taught, but not in the context of the trebuchet. That Laura was able to apply the lever concept to her trebuchet shows subject matter learning. Laura's think aloud excerpt for this example also shows subject matter learning as well as metacognition.

> A trebuchet is best represented by a series of simple machines that come together to form one common machine at school. For example, the main machine is located at the first-class lever, that the… Tricky (M) should be force of it. The projectiles output force out of its reach, the closer the load is to the fulcrum. The more this lever can thus.... The more they lever potentially launch it okay I'm gonna finish now (P).

Both the writing and the think aloud indicate that Laura can identify a first class lever and understand the relationship between the fulcrum and the output force. Laura also engaged in monitoring and planning as she moved through this example.

This was one example in an overall more sophisticated scientific explanation. As a more complex understanding of the subject matter is shown, a higher corresponding TA score is also present. Again this example shows subject matter learning from argumentation and metacognition.

Summary of Results by Research Question

The ACORN framework can be used to produce written argumentation, does promote metacognition, and gives evidence of student learning. Students were able to consistently show the elements of claims, evidence, and reasoning as a part of their written responses therefore demonstrating argumentation (Bricker & Bell, 2008). Students also consistently engaged in regulation of cognition, one component of metacognition. While engaging in metacognition during argumentation was expected (Erduran & Jiménez-Aleixandre, 2008; Garcia-Mila, & Andersen, 2007; Zohar, 2007), qualitative data showed that the use of the ACORN framework promoted regulation of cognition. Subject matter learning was also shown qualitatively in the student responses, student interviews, and think alouds. More importantly, links among argumentation, metacognition, and subject matter learning were shown.

CHAPTER 5

DISCUSSION AND IMPLICATIONS

Conceptual Framework Revisited

The conceptual framework for this study is based on four pillars supporting metacognition which in turn supports student subject matter learning. The first pillar, coordination, is a scientific process that requires students to manipulate information. In this research, coordination happened when students needed to cite information and provide reasoning as part of their generating a written argument. Students needed to decide which pieces of evidence from a wide variety sources best supported their argument. Then students continued by providing reasoning to show how their evidence supported to their claim. The coordination of evidence and reasoning to support the claim is fostered by the Organize, Respond, and Note steps of the ACORN framework. Qualitative analysis of the think alouds shows that this coordination supports student metacognition. First, the nine students involved with the Think alouds directly referred to the ACORN process 41 times. Second, these students had a total of 374 instances of metacognition. The mentions of ACORN coinciding with instances of metacognition show how the ACORN framework is part of the process of coordinating information to elicit metacognition. Laura states, “okay, let me start with my respond (A)(P). There’s a better way to write it. (M)”. The letters in the parenthesis indicate instances of metacognition where (P) indicates procedural regulation of cognition and (M) indicates monitoring regulation of cognition while the (A) indicates a direct reference to the ACORN framework. Laura is directly referring to the ACORN framework's Respond while coordinating information resulting in instances of regulation of metacognition planning and monitoring.

The second pillar of the conceptual framework is evidence and is both a scientific process and scientific content. Here students must decide what counts as evidence (process related) and what evidence to use (content related). This is supported in the ACORN framework through the Cite step. For example, in the passage from one of the think alouds cited below, Erica reflected on the activities and content in class to decide which pieces of evidence can be used to support a claim. She demonstrates the process of weighing evidence to support a claim. The letters in the parenthesis indicate instances of metacognition where (P) indicates procedural regulation of cognition and (M) indicates monitoring regulation of cognition.

> Okay, so, what are the choices? (P) Uhh, I think I will do either (P) second law, or simple machines because it is the throwing arm of a lever, the throwing arm of the trebuchet is a lever. I need to figure out which (P) one would be best. Umm, I mean I should probably do simple machines since I know its parts (M), like fulcrum and load and output force, I think that's it. (M) Umm, Newton's second law is masses equals force times acceleration. How would that work with the hmm,(M) umm so so I guess I could say that the projectile and a load had the same force

All but one of the student think alouds demonstrated similar instances of metacognition when deciding on evidence which also shows that evidence supports metacognition.

The third and fourth pillars of the conceptual framework are prior and new knowledge. These pillars do not directly produce instances of metacognition, but represent the content needed to demonstrate that subject matter learning has taken place. Even though these pillars do not directly produce metacognition, it is attention to these pillars that allow for metacognition in the context of argumentation. Being able to link prior knowledge to new knowledge – link evidence with reasoning to support a claim - and to assess that link is where metacognition is produced. For example, in a Think Aloud, Enid debates whether her evidence (prior knowledge) of Newton's Laws is enough to support her claim (new knowledge). She states,

> I am going to write about the third which was was about forces acting in pairs. Now I remember. (M) I'm good write about Newton's third law which is so much better because force action pairs, (M) so when you pull back, it goes. Okay I'm going to write about Newton's third law. (P)

Knowing that forces act in pairs and that can be represented by pulling back on the trebuchet qualifies as her evidence (prior knowledge). The application of her evidence to describe the motion of the trebuchet supports her eventual deeper understanding of Newton's third law (new knowledge). Here we also see instances of metacognition in that she is using prior knowledge and new knowledge. In the context of argumentation, these pillars need to be present in order for coordination to take place. In other words without scientific content – prior and new – there would not be scientific argumentation. Without scientific argumentation, metacognition will not be produced in order to support student subject matter learning.

The importance of the four pillars in the conceptual framework is evident in the results of this research. Furthermore, the ACORN framework was developed specifically to provide opportunities for metacognition during the argumentation process. Students need to be taught argumentation and metacognitive skills (Garcia-Mila, & Andersen, 2007; Kuhn & Udell, 2003). This research shows the promise introducing metacognition into written argumentation as means to promote student subject matter learning (Bangert-Drowns, Hurley, & Wilkinson, 2004).

Discussion

Research Question 1

The purpose of the first research question is to map argumentation onto the ACORN framework. When scoring each paper, claims, evidence, and reasoning can be seen in the students writing. Specifically the Answer step of ACORN relates directly to the claim, the Cite step of ACORN relates to the evidence and the Organize and Respond steps relate directly to the reasoning. In table 5.1, the claims, evidences, and reasoning of the lowest and highest scoring final writings are given.

Table 5.1. Examples of Claims, Evidence and Reasoning

Argumentation element	High scoring	Low scoring
Claim	Simple machines best represent the workings of the trebuchet	A simple machine is best represented in a trebuchet.
Evidence	Simple machine = an object used to demonstrate same amount of work done with varying forces depending on the type	A lever is one of the simple machines inside of a trebuchet
Reasoning	Wheels allow arm to swing in a straighter and not lose energy so the load goes further. Also keeps the trebuchet stable and balance w/ the recoil.	The throwing arm would be the lever and the metal bar is the fulcrum.

Looking at both ends of the spectrum, claims, evidence and reasoning are present in both. It is important to note that these examples were pulled directly from the ACORN graphic organizer as seen in Figure 5.1

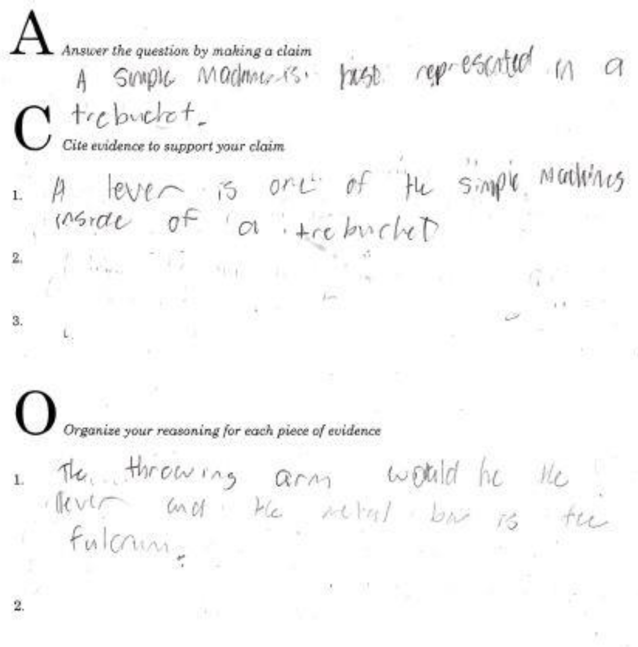


Figure 5.1. Student Graphic Organizer Showing ACORN Process.

Using the ACORN process and associated graphic organizer provides the structure needed for a student to present claims, evidence and reasoning thus promoting written argumentation.

Throughout the intervention, students were able to improve their written argumentation using the ACORN framework. This is important because being able to show improvement speaks to student learning of argumentation skills. If the ACORN framework did not allow for student argumentation, then the students would not be able to improve on their argumentation. Students were able to consistently improve their writing, according the TA rubric, with the exception of the Simple Machines writing. There are several contributing factors to explain the lower scores on the Simple Machine writing. First, this assignment was administered differently from the other four writings. For this writing, students completed this as a constructed response as a part of a simple machines test as opposed to a unique assignment as the other writings were. This caused the students to have less time to complete this writing as opposed to the others. Second, the question itself was designed to demonstrate the students' knowledge of levers and did not present an option of arguing different sides of the issue. While a student could use different evidences and reasonings to answer the question, the question itself locked students into one claim. Not having the ability to argue different claims changes the nature of written argumentation compared to the other assignments where students had more freedom to select a claim to defend.

Students were also able to consistently improve their writing according to the HA rubric. Again, students scored lower on the simple machines writing likely for the same reasons noted above. Also, the HA scores on the SVA writing were high. In retrospect this also makes sense. With the SVA writing, there was one clear correct answer to the question thus making the writing task easier to complete. Students did not need to contemplate the possibilities of multiple correct solutions and could focus on the facts of matter. Also, being able to read a time distance graph in order to complete the assignment was easy for the participants. That the HA scores were higher on an easier assignment lends credence to the idea that students can effectively use ACORN to produce written argumentation. One would expect higher argumentation scores when the case is easier to argue. It is also interesting to note that the TA scores for the SVA writing were not higher as well. In other words, students were able to holistically show their increased argumentation ability on an easier assignment, but not necessary show same increase in structure of argumentation. This shows the importance of using two different rubrics to assess the students' writings. This also raises questions for future research in the differences between a structural and holistic approach to written argumentation.

Each of the essential elements of argumentation - forming a claim, supplying evidence, and using reasoning to link the evidence to the claim – is a part of argumentation (Bricker & Bell, 2008), but understanding how these elements work together is the real power behind argumentation (Bricker & Bell, 2008; Duschl, 2007). The ACORN framework allows for these elements and for the relationship among these elements since the normal distribution of scores on both TA and HA rubrics as well as the improvement in writing can be shown. Furthermore the mapping of claims, evidence, and reasoning as a result of using ACORN can be shown. Showing that ACORN promotes argumentation is an important step in understanding how students' metacognition during argumentation promotes student subject matter learning. Several frameworks exist to promote argumentation such as CER, SWH, and ADI. On one end of the spectrum, CER address the common elements present in argumentation claims, evidence, and reasoning (McNeill & Martin, 2011), and is possibly as a straight forward approach as it gets. If CER is such a direct approach, why do other frameworks exist such as SWH and ADI? These two frameworks allow for inquiry driven activities to be dissected in order to produce written argumentation. They are sharp tools that when used correct can produce the desired results (Wallace, Hand, & Prain, 2004; Grooms, Enderle, & Sampson, 2015), but they are designed specifically to be used with inquiry based instruction (Wallace, Hand, & Prain, 2004; Grooms, Enderle, & Sampson, 2015). CER on the other hand is easy, straight forward, but it also blunt. It does not allow for the sophistication needed to promote metacognition and thus subject matter learning on a meaningful level. ACORN has attempted to be the in between instrument that can be used in a wide variety of situations, yet sophisticated enough to produce metacognition in students as addressed by research question number two.

Research Question 2

The ACORN framework was designed to promote metacognition in the writing process that would lead to subject matter learning (Bangert-Drowns, Hurley, & Wilkinson, 2004). Specifically, the Organize step and the Note step were inserted to give the students the opportunity to contemplate their own writing. The purpose of the Organize step is to allow students to organize the reasoning behind the evidence they cited. Organizing involves monitoring because it requires the students to think about what they are doing with their reasoning and how it relates to the evidence, not just providing an explanation. Figure 5.2 shows the Cite and Organize step for Nancy's final writing.

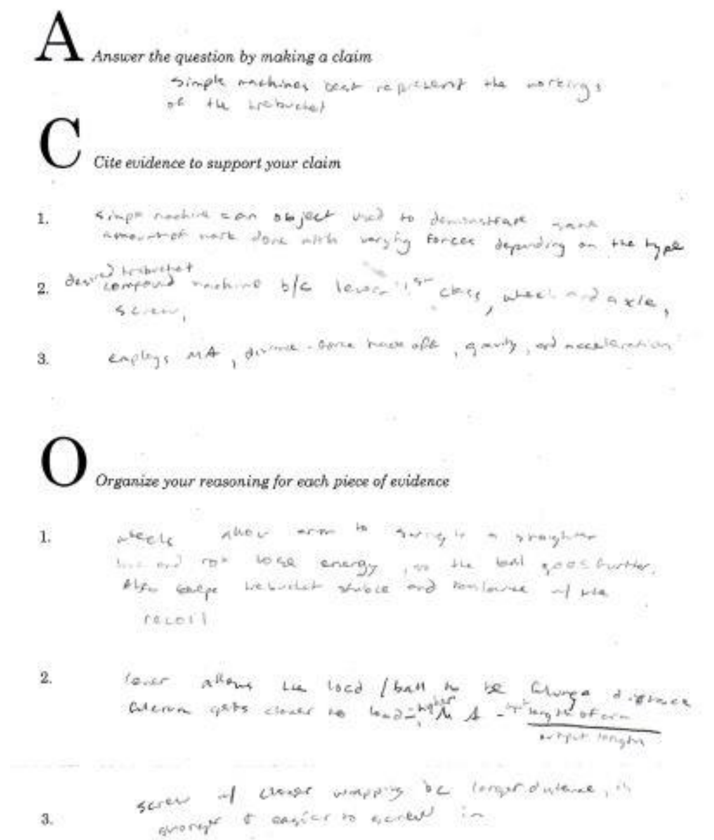


Figure 5.2. Nancy's Final Writing Graphic Organizer

Here is the associated dialogue with the second bullet point under the Organize section of the graphic organizer:

> So basically, I'm going to break down this trebuchet and talk about each of its inner workings and basically what the video said I, I (M) was and what kind of plan actually, I think it is so much better when I have more time to plan (PK), but that's what everything, I do everything so slowly (DK) ummm and today a social studies test great. Ooooh let's talk about the lever next (P) . The lever allows the load/ball to be flown a distance. As fulcrum gets closer to load, mechanical advantage or lenght of arm divided by like the input/ouput length. Pretty sure that's it. (M) Equals higher mechanical advantage, right? (M) Yes.

Matching wording in the dialog aligns the Think Aloud with the Organize step. In this one step, six instances of metacognition are present. Furthermore, in the interview Nancy states, "the Cite and Organize helped me the most." The organize step promotes metacognition because it requires the students to think about how their reasoning relates to the evidence.

The Note step directly asks the student to engage in metacognition. The purpose of the Note is for students to be asked to reflect on their work in some manner. For the final writing, students were asked to discuss what helped them most in completing their writing. This statement is an evaluation of their writing and their experiences purposely designed to elicit metacognition. Laura's Note for her final written was "Thinking back to my group's trebuchet experience helped me complete my writing because I was able to use an actual class/real-life experience to relate my thoughts back to. Organizing my ideas also helped me piece everything together." Not only did this Note provide an instance of evaluation, but also adds evidence that the Organize step in the ACORN process is valuable to students.

One of the unexpected results from this research was that there was a weak correlation between the MAI post-test regulation of cognition scores to both the TA score for the final writing (r=0.058) and the HA score for the final writing (r=.15). Since the link between metacognition and writing has been established (Flower & Hayes, 1981; Mayher & Lester, 1983; Applebee, 1984, Langer & Applebee, 1987; Bereiter & Scardamalia, 2013) and the link between metacognition and argumentation has been established (Kuhn, 1992), it would be reasonable to expect a correlation between metacognition as measured by the MAI and the argumentation scores of the final writing. Some possibilities exist though to explain this lack of correlation. First and foremost is the ability of a 13- or 14-year-old student to accurately assess their own metacognitive awareness. "The hypothesis that thinking about one's own thought does not emerge until late childhood or early adolescence appears to be correct" (Kuhn, 1992, p. 172) although young children do display the awareness of "their own and others mental functions"

(Kuhn, 2000, p.180). These two ideas present by Kuhn place early adolescents in a gray area of whether or not they possess metacognitive processes much less whether they are aware of them. If these students are at an age where metacognitive ability is developing, at least some of the students could possibly be unable to accurately report on their own metacognition.

Second, the MAI was initially developed with undergraduate students (Schraw and Dennison, 1994). The authors state that the MAI provides a "reliable initial test of metacognitive awareness among older students" (Schraw and Dennison, 1994, p. 472), but no qualifying age was given. It is possible that the MAI is not suited for middle school aged students, however the high instances of metacognition of regulation found in the student think alouds with the significant increase in the MAI regulation score provides at least some evidence to the contrary. Furthermore, considering the performance level of students, both HS (high scoring) students and HM (high metacognitive) students displayed high average numbers of regulation of cognition compared to the LSLM students in the think alouds. Convergence of the data leads me to believe that the MAI was a reasonable choice for this research, but future research might include different instruments for measuring metacognitive awareness.

A third possible explanation involves a deeper examination of the development of metacognitive skills and how it relates to argumentation. Kuhn (1992, 2000) suggests that it is the social construction of dialogic argumentation that allows for development of metacognitive thinking. In other words, students must be able to converse with other students in order for argumentation to have benefit to metacognition. Graff (2003) agrees suggesting that argumentation is difficult for students and that being involved in dialogue will ultimately improve written argumentation. It is interesting to note here that both the SWH and ADI frameworks allow for this dialogue where it is not an explicit step in CER, RACE, or ACORN.

Since students did not have the opportunity to engage in dialogue when developing their argument, and dialogue promotes metacognition in argumentation, then it is possible this would result non-significant correlation between TA and HA score and the MAI. In essence, the lack of dialogue is the missing link between high scoring written argumentation and metacognition. If this is the case, then future research would need to examine if adding a social construction piece to ACORN would affect metacognition, scores, and the relationship between them.

Overall this research shows that students do engage in metacognition using the ACORN framework through quantitative results of the MAI, analysis of student think alouds, and student interviews. This research also did not show a link between metacognition and argumentation scores as measured by the TA and HA rubrics. The question then becomes whether or not introducing metacognition is enough of a benefit of using the ACORN framework. The answer to this question seems to lie in whether or not the ACORN framework promotes subject matter learning.

Research Question 3

The purpose of research question 3 was to connect the student performance in argumentation to student subject matter learning. There was a significant improvement from FCI pretest scores and FCI post test scores ($p=1.7\times10^{-20}$ and $\alpha=0.05$) showing that there was student subject matter learning during the time of the intervention, but there was a weak correlation between the FCI post test scores and the TA ($r=-.02$) or HA ($r=-.13$) final writing scores. Without the link between the argumentation TA and HA scores and the performance on the FCI post-test, other possibilities for the weak correlations must be considered.

First, the FCI did not measure the concepts of simple machines, work and power, yet 18 out of 48 of the final responses were about simple machines, work, and power. This creates a situation where 37.5% of the students are writing about something that is not measured by the FCI. In hindsight, the option of simple machines as a topic for the final writing should not have been available to students. For sake of argument though, a linear regression was performed between the FCI post-test scores and final argumentation scores with the students who responded about simple machines removed. This resulted in 30 students who wrote on their final assignment about topics measured by the FCI. Again only weak correlations existed with the TA score resulting in r=0.32 and the HA score resulting in r=0.04. This suggests that a reason for weak correlation between argumentation scores and the FCI post-test scores lies elsewhere.

Another possibility to explain the discrepancy between argumentation scores and the FCI post-test scores has to do with the nature of the middle school student and the length of the unit involved. This unit spanned nine weeks and covered many topics. The students were not offered a review before the administration of the FCI posttest. It is likely that a structured review time, as appropriate for young adolescents (Manning, 1993), would have impacted the results of the FCI. On the contrary, the absence of a review might be a better demonstration of what the students may have learned at a deep level.

The third possibility is that students learned the content as a result of the other activities in class not related to argumentation. This is the most likely of reasons since a variety of learning activities were concurrent with writing using ACORN. It would be impossible to use argumentation as a single instructional strategy for a long unit, and in order to see the effects of ACORN a quasi-experimental design would need to be implemented. It is difficult to concede that ACORN did not promote student subject matter learning while other activities did since the writing was intertwined with classroom instruction. Also, student subject matter learning can be seen in the student artifacts as well as in the student interviews.

Student interviews indicate that students felt like they learned. Repeatedly, the students who were interviewed affirmed that writing using ACORN helped them make sense of the information presented in class. This corroborates the idea that writing can promote student learning (Emig, 1977; Flower & Hayes, 1981; Mayher & Lester, 1983; Applebee, 1984, Langer & Applebee, 1987; Keys, 1999; Bereiter & Scardamalia, 2013). Since the FCI scores did not strongly correlate to the argumentation scores, but student interviews did indicate subject matter learning, a different measure of subject matter learning should be investigated.

The primary source demonstrating student subject matter learning is the scientific explanations the students have written. The TA rubric relies on the strengths of the claims, evidences, and reasonings by definition. In order to score high on the TA rubric, students must demonstrate quality content. The categories of quality of evidence, relationship between claims and evidence, and reasoning all relate to the content learned in class. In other words, it would be difficult for a student to score well on the TA rubric if the student had not learned the content. For example, a weak score in the quality of evidence category suggests a student has copied information from the textbook, described data, or presents inaccurate evidence. On the contrary a very high score in the quality of evidence category suggests that a student has very accurate, credible, reliable evidence interpreted from observations and data. Since the criteria for a high-scoring writing on the TA rubric relies on content, the TA score can be used as a proxy for student subject matter learning. If this is the case, a comparison of the TA score as a measure of subject matter learning, and the HA score as a measure of argumentation might prove interesting. In fact the correlation between the TA Final score and the HA Final score is $r=.76$. This would show that holistic argument is strongly correlated to subject matter learning as reflected in the TA score. While this is not a profound measure of student subject matter learning as a result of argumentation, it does provide some evidence to support subject matter learning through argumentation.

Any marker used to represent student subject matter learning has the possibility of ignoring different elements of learning. That the FCI does not correlate to the TA and HA final scores does not show that subject matter learning as a result of written argumentation did not occur. Qualitative evidence through the student writings and student interviews does suggest that such subject matter learning takes place, indicating that further research is necessary to a better establish the link between the ACORN framework and student subject matter learning.

Implications

The overarching goals of this research was to establish the ACORN framework as a legitimate tool for students to produce written argumentation, to have students engage in more metacognitive activities while producing written argumentation, and to show that ACORN can be used to promote student subject matter learning. In achieving these goals, I wanted to give students and teachers a powerful instructional tool.

ACORN is clearly a tool that can be used for argumentation. It helped student to organize thoughts, express their ideas using the tenets of argumentation, and structure a scientific explanation. Student interviews indicated that there was initial confusion about the acronym, but that they became more comfortable with the meaning of each step as they gained experience. While this might be expected learning any new concept, my hope was that the acronym would be readily self explanatory. In future iterations of this research, an adjustment to the acronym might be considered. More likely though, more time and examples might be given when introducing each step of the framework to the students. One effective way of teaching students writing is to model that process through use of examples (Krashen & Terrell, 1995). Krashen & Terrell

(1995) outline an approach for language acquisition. They suggest that instead of concentrating on the syntax of the language, that the teacher needs to concentrate on the message given. This is primarily achieved through examples whether oral or written. Considering that argumentation is considered a discourse in the scientific community, it acts as a unique language. It is reasonable to assume that written examples would help students learn argumentation. Having never used ACORN before, there was no work available to show students for the first writing. Had there been an anchor paper and sample graphic organizers to use, I feel that many questions about the steps in ACORN could have been addressed in the beginning.

One of the important results to come from ACORN in the sense of written argumentation was the ability of the students help distinguish between evidence and reasoning. This is a difficult task (Sampson & Clark, 2008, Sandoval & Millwood, 2007) and that students indicated a better ability to do this is encouraging. Being able to differentiate between evidence and reasoning takes the scientific process of coordination which in turn promotes learning. I have seen many students struggle with this concept and giving them a tool to separate evidence from reasoning aids in students learning not just in science content, but in science processes.

Another foundational aspect of this research was to introduce an avenue for metacognition when writing scientific explanations. The ACORN framework does this as well. While most people agree that metacognition is a good for subject matter learning, I have seen plenty of teachers not include the use of metacognition in their lessons. Since the ACORN framework explicitly allows for metacognition in the Organize and Note steps, the use of the framework promotes metacognition skills in the classroom as a part of instruction. This is invaluable to teachers and student in order to promote student learning.

It was discouraging that the results of the MAI did not correlate to the TA and HA scores. Metacognition is difficult to assess though (Schraw, Crippen, & Hartley, 2006) and future iterations of this research need to be mindful of that. The difficulty in measuring metacognition was one reason for the mixed methods study because I felt that metacognition could best be assessed qualitatively. The weak correlation between writing scores and the MAI does not show that students did not engage metacognitively. All it suggests is that a different instrument be used to gauge metacognition since metacognition was demonstrated through the think alouds and the student interviews. This is particularly highlighted when comparing the HSLM performance level students to the number of instances of metacognition on the student think alouds. HSLM students actually had the highest average instances of regulation of cognition compared to the other three performance levels. While there is not enough quantitative data for a solid conclusion, the trend seems to indicate that the MAI might not accurately capture metacognition. In future research, I would have a tendency to rely on instances of metacognition as shown in the think alouds instead of a quantitative measure of metacognitive self awareness.

Also disappointing was the lack of correlation between the student writings and the score on the FCI. Given the low average scores on the FCI though, a better measure of student subject matter learning might be appropriate. First, I would include more frequent assessments of subject matter learning so that students would not have nine weeks of material on which to be assessed. I would also have the assessments be more middle school friendly.

The next step for the ACORN framework would be to set up a quasi-experimental research designed to show the differences in student subject matter learning with ACORN as the intervention. The primary reason for this stems from the significance in subject matter learning as shown by the FCI yet the lack of correlation between student subject matter learning and the writing scores. Also further research into student learning and metacognition could be performed by multiple case studies involving student writing and think alouds. This could examine the relationship ACORN has to metacognition and student learning.

Another direction for research would be to investigate argumentation as student learning. The importance for this research is highlighted in the NGSS. Argumentation is specifically listed in the NGSS as a learning goal (Lead States, 2013). Additionally, more credibility is being given to so called 21st century skills that emphasis how students process information as opposed to the memorization of information. While not specially examined as a part of this research, argumentation learning would highlight argumentation as important student learning goal. This research does however provide a glimpse into argumentation learning. Quantitatively, the writings improved according to the TA and HA rubrics indicating a form of student learning. In other words, this shows that using the ACORN framework that students were able to learn the elements of argumentation.   Qualitatively, the ACORN framework was seen as an organizational tool by the students who used it. It is possible that this organization leads to argumentation learning.   An analysis of this data from the perspective of student argumentation learning would be interesting and could possibly contribute to the argumentation literature as a whole. Such an analysis was left out of this research because it was outside the scope of the research questions as this research was founded in the context of argumentation as a pedagogical tool for subject matter learning.

Ultimately, ACORN is a tool to promote written argumentation and metacognition in the classroom. Student subject matter learning as a result of ACORN has yet to be shown. Pragmatism holds that the results of research are not final truths (Johnson & Onwuegbuzie, 2004). Similarly, science is constantly changing and adapting based on new evidence

(Lederman, 2007). Future research will show whether ACORN will be an effective tool to promote written argumentation, metacognition, and student subject matter learning.

REFERENCES

Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., & Tuan, H. L. (2004). Inquiry in science education: International perspectives. *Science Education*, *88*, 397-419.

Applebee, A. N. (1984). Writing and reasoning. *Review of Educational Research*, *54*, 577-596.

Baker, L., & Cerro, L. C. (2000). Assessing metacognition in children and adults. In G. Schraw & J. Impara (Eds.) *Issues in measuring metacognition* (pp. 99-145). Accessed 1/24/2017 from http://digitalcommons.unl.edu/classicsfacpub/62/

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, *74*, 29-58.

Bereiter, C., & Scardamalia, M. (2013). *The psychology of written composition*. (Kindle Fire edition). Routledge.

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*, 765-793.

Biesta, G., & Burbules, N. C. (2003). *Pragmatism and educational research*. Lanham, MD: Rowman & Littlefield Publishers.

Bricker, L. A., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education*, *92*, 473-498.

Choi, A., Notebaert, A., Diaz, J., & Hand, B. (2010). Examining arguments generated by year 5, 7, and 10 students in science classrooms. *Research in Science Education*, *40*, 149-169.

Creswell, J. W. (2014). *A concise introduction to mixed methods research*. Thousand Oaks, CA: Sage Publications.

Denzin, N. K. (1996). Post-pragmatism. *Symbolic Interaction*, *19*, 61-75.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*, 287-312.

Duschl, R. A. (2007). Quality argumentation and epistemic criteria. *Argumentation in science education*, 159-175.

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education, 38,* 39-72.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Academies Press.

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication*, *28* 122-128.

Erduran, S., & Jiménez-Aleixandre, M. P. (2008). *Argumentation in science education: Perspectives from classroom-based research.* Dordrecht, Netherlands: Springer.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*, 365-387.

Garcia-Mila, M., & Andersen, C. (2007). Cognitive foundations of learning argumentation. In S. Erduran & M.P. Jimenez-Aleixandre (Eds.) *Argumentation in science education* (pp. 29-45). New York, NY: Springer.

Glaser, B. (1965). The constant comparative method of qualitative analysis. *Social Problems, 12,* 436-445. doi:10.2307/798843

Graff, G. (2003). Clueless in *academe: How schooling obscures the life of the mind*. New Haven, CT: Yale University Press.

Greene, J. C. (2007). *Mixed methods in social inquiry* (Vol. 9). San Francisco, CA: John Wiley & Sons.

Grooms, J., Enderle, P., & Sampson, V. (2015). Coordinating scientific argumentation and the next generation science standards through argument driven inquiry. *Science Educator*, *24*, 45.

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice*. New York, NY: Guilford Press.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*, 141-158.

Himmele, P., Himmele, W., & Potter, K. (2014). *Total literacy techniques: Tools to help students analyze literature and informational texts*. Alexandria, VA: ASCD.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14-26.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*(2), 112-133.

Kellough, R. D., & Kellough, N. G. (1999). *Middle school teaching: A guide to methods and resources*. Upper Saddle River, NJ Merrill.

Keys, C. W. (1999). Revitalizing instruction in scientific genres: Connecting knowledge production with writing to learn in science. *Science Education*, *83*, 115-130.

Kiefer, K. (n.d.). The WAC Clearinghouse. Retrieved April 27, 2015, from http://wac.colostate.edu/intro/pop2d.cfm

Krashen, S. D., & Terrell, T. D. (1995). The natural approach: Language acquisition in the classroom. Cornwall, Great Britain, Prentice Hall Europe.

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review*, *62*, 155-179.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*, 178-181.

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, *22*, 545-552.

Kuhn, D., & Dean, Jr, D. (2004). Metacognition: A bridge between cognitive psychology and educational practice. *Theory into Practice*, *43*, 268-273.

Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development*, *74*, 1245-1260.

Langer, J.A., & Applebee, A. (1987). *How writing shapes thinking:* A *study of teaching and learning* (NCTE Research Rep. No. 22). Urbana, IL: National Council of Teachers of English.

Lederman, N. G. (2007). Nature of science: Past, present, and future. In N. Lederman & S. Abell (Eds.), *Handbook of research on science education*, (vol. 2, pp. 831-879).

Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to Next Generation Science Standards and with implications for Common Core State Standards for English language arts and mathematics. *Educational Researcher, 42(4)*, 223-233. doi: 0013189X13480524.

Lead States. (2013). *Next generation science standards: For states, by states*: Appendix F

Mayher, J. S., & Lester, N. B. (1983). Putting learning first in writing to learn. *Language Arts*, *60*, 717-722.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences*, *15*, 153-191.

McNeill, K. L. & Martin, D. M. (2011). Claims, evidence and reasoning: Demystifying data during a unit on simple machines. *Science and Children. 48(8)*, 52-56.

Nam, J., Choi, A., & Hand, B. (2011). Implementation of the science writing heuristic (SWH) approach in 8th grade science classrooms. *International Journal of Science and Mathematics Education*, *9*, 1111-1133.

O'Cathain, A. (2010). Assessing the quality of mixed methods research:Toward a conceptual framework. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (2nd ed., pp. 531-555).

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, *41*, 994-1020.

Pauli, D., (2017, April). *The RACE framework as written argumentation in 8th grade physical science*, Poster presented at the annual meeting of the National Association for Research in Science Teaching conference, San Antonio, TX.

Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. Impara (Eds.) *Issues in measuring metacognition* (pp. 99-145). Accessed 1/24/2017 from http://digitalcommons.unl.edu/classicsfacpub/62/

Sampson, V., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education*, *92*, 447-472.

Sampson, V., Enderle, P., Grooms, J., & Witte, S. (2013). Writing to learn by learning to write during the school science laboratory: Helping middle and high school students develop argumentative writing skills as they learn core ideas. *Science Education*, *97*, 643-670.

Sampson, V., & Gleim, L. (2009). Argument-driven inquiry to promote the understanding of important concepts & practices in biology. *The American Biology Teacher*, *71*, 465-472.

Sandoval, W. A., & Millwood, K. A. (2007). What can argumentation tell us about epistemology? . In S. Erduran & Jimenez-Aleixandre (Eds.) *Argumentation in science education* (pp. 71-88). Dordrecht, Netherlands: Springer.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, *26*, 113-125.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460-475.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, *36*, 111-139.

Street, C. (2003). Pre-service teachers' attitudes about writing and learning to teach writing: Implications for teacher educators. *Teacher Education Quarterly*, *30*(3), 33-50.

Tashakkori, A., & Creswell, J. W. (2007). Editorial: Exploring the nature of research questions in mixed methods research. *Journal of Mixed Methods Research*, *1*, 207-211.

Toulmin, S. E. (2003). *The uses of argument*. New York, NY: Cambridge University Press.

Van Eemeren, F. H., Grootendorst, R., Henkemans, F. S., Blair, J. A., Johnson, R. H., Krabbe, E. C., ... & Zarefsky, D. (1996). *Fundamentals of argumentation theory*. New York, NY: Routledge.

Von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, *45*, 101-131.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT press.

Wallace, C. S., Hand, B. B., & Prain, V. (2004). *Writing and learning in the science classroom* (Vol. 23). New York, NY: Springer Science & Business Media.

Zohar, A.. (2007). Science teacher education and professional development in argumentation. In S. Erduran & M.P. Jimenez-Aleixandre (Eds.) *Argumentation in science education* (pp. 29-45). New York, NY: Springer.

APPENDIX

Student Interview Questions

The following questions were a starting point for the final version of the interview questions. The results of the initial quantitative analysis were used to help inform the development of the question, but these questions or variations on these questions were included.

Student Interview Questions

1. What did you think about writing using ACORN? What do you mean by that?
2. What do you feel were the most beneficial about using ACORN? Why?
3. What do you not like about using ACORN paragraphs?
4. Explain how you think you learned from this process.
5. Explain how you think the process could be improved.
6. What did you learn?
7. Do you think that the ACORN framework related to the activities we did in class? Why or why not?
8. Explain how you thought about learning as you went through this process.
9. Here is a copy of your writing. Walk me through how you wrote this?
    a. What part was the easiest to write? Why?
    b. What part was the hardest to write? Why?
    c. What did you mean by this?
    d. Do you think this helped you learn anything? Why or why not?
10. What else would you like to tell me about this unit?

Metacognition Awareness Inventory

The following is the list of 52 questions included on the MAI. For further details, see Schraw & Dennison (1994).

1. I ask myself periodically if I am meeting my goals.

2. I consider several alternatives to a problem before I answer.

3. I try to use strategies that have worked in the past.

4. I pace myself while learning in order to have enough time.

5. I understand my intellectual strengths and weaknesses.

6. I think about what I really need to learn before I begin a task

7. I know how well I did once I finish a test.

8. I set specific goals before I begin a task.

9. I slow down when I encounter important information.

10. I know what kind of information is most important to learn.

11. I ask myself if I have considered all options when solving a problem.

12. I am good at organizing information.

13. I consciously focus my attention on important information.

14. I have a specific purpose for each strategy I use.

15. I learn best when I know something about the topic.

16. I know what the teacher expects me to learn.

17. I am good at remembering information.

18. I use different learning strategies depending on the situation.

19. I ask myself if there was an easier way to do things after I finish a task.

20. I have control over how well I learn.

21. I periodically review to help me understand important relationships.

22. I ask myself questions about the material before I begin.

23. I think of several ways to solve a problem and choose the best one.

24. I summarize what I've learned after I finish.

25. I ask others for help when I don't understand something.

26. I can motivate myself to learn when I need to

27. I am aware of what strategies I use when I study.

28. I find myself analyzing the usefulness of strategies while I study.

29. I use my intellectual strengths to compensate for my weaknesses.

30. I focus on the meaning and significance of new information.

31. I create my own examples to make information more meaningful.

32. I am a good judge of how well I understand something.

33. I find myself using helpful learning strategies automatically.

34. I find myself pausing regularly to check my comprehension.

35. I know when each strategy I use will be most effective.

36. I ask myself how well I accomplish my goals once I'm finished.

37. I draw pictures or diagrams to help me understand while learning.

38. I ask myself if I have considered all options after I solve a problem.

39. I try to translate new information into my own words.

40. I change strategies when I fail to understand.

41. I use the organizational structure of the text to help me learn.

42. I read instructions carefully before I begin a task.

43. I ask myself if what I'm reading is related to what I already know.

44. I reevaluate my assumptions when I get confused.

45. I organize my time to best accomplish my goals.

46. I learn more when I am interested in the topic.

47. I try to break studying down into smaller steps.

48. I focus on overall meaning rather than specifics.

49. I ask myself questions about how well I am doing while I am learning something new.

50. I ask myself if I learned as much as I could have once I finish a task.

51. I stop and go back over new information that is not clear.

52. I stop and reread when I get confused.

Total Argument and Holistic Argument Rubrics

Scoring Rubric for the Total Argument Score

	Quality of Questions	**Quality of Claims**	**Relationship Between Questions and Claims**	**Quality of Evidence**	**Relationship Between Claims and Evidence**	**Reasoning**
1	Single question Closed-ended question Questions are not testable Unimportant and poor questions Questions does not catch the essence of the investigation Questions are insignificant Questions are of low quality	Single Claim Claims are not based on any data or observation Claim does not catch the essence of the investigation Claim is insignificant Claim is invalid and inaccurate Claim is of low quality	Very weak connection between questions and claims Claims without any questions or questions without any claims Questions and claims do not fit at all	Very weak evidence, inaccurate, invalid, and unreliable evidence Evidence is very sparse Their observation is itself evidence (e.g., see my observation, calculation, or data section) Evidence seems to come form no where in particular	Very weak connection between claims and evidence Evidence is not focusing on the claims at all Claims without evidence or evidence without claims	Very weak explanation of how the evidence relates to the claim Student is not able to link to their own investigation to their existing knowledge Student does not spot errors
2	A few questions Closed-ended questions Testable or may be difficult to test May not be meaningful questions	Single or multiple claims Claims may not appear to have come from their experimental observation/data Claims may not	Weak connection Questions and claims fit loosely Student develops claims for a few of the generated questions	Weak evidence May not be accurate, valid, and reliable Evidence are just a description of data Evidence are from	Weak connection Evidence supports claims loosely or inadequately Students provide evidence for a few claims	Weak explanation for how the evidence relates to the claim Student may not be able to link their own investigation to their existing knowledge Student may not spot

	Questions may not catch the essence of the investigation Questions may not be significant and adequate Questions may be of low quality	catch the essence of the investigation Claims may not be significant and adequate Claims may not be valid and sound Claims may be of low quality	Claims are uncertain in answering questions	textbook	Proposed evidence may not be apparent in supporting claims Evidence is focusing on a few claims loosely	errors
3	Multiple questions which are primarily closed-ended questions If only one, it is open-ended question. Testable and meaningful questions Questions may match the essence of the investigation Questions may be significant and adequate Questions may be of high quality	Single or multiple claims Claims may be from their experimental observation/data Claims may match the essence of the investigation Claims may be significant and adequate Claims may be valid and sound Claims may be of high quality	Moderate connection Questions and claims fit reasonably Student develops claims for some of the generated questions and the proposed claims may be apparent in answering questions. Claims are focusing on all the questions but loosely connected with the questions	Moderate evidence May be valid evidence May be reliable evidence Evidence from textbook with a limited interpretation or explanation	Moderate connection Evidence supports claims reasonably Students provide evidence for some of the generated claims and proposed evidence may be apparent in supporting claims. Evidence is focusing on all the claims but loosely connected with claims.	Moderate understanding of how the evidence relates to the claim Student may understand how their investigations tie into concepts about what they have learned in class Student may make connections to concepts Student may spot errors and may not explain them
4	Multiple questions which include at least one open-ended question Testable questions Questions catch the essence of the investigation Questions are significant and adequate Questions are of high quality	Multiple claims Claims are from the interpretation of their experimental observation/data Claims catch the essence of the investigation Claims be significant and adequate Claims be valid, sound, and accurate Claims be of high quality	Strong Connection Questions and claims fit strongly Student develops claims for most of the generated questions Proposed claims are evident in answering questions even though claims are only for some of the generated questions	Powerful evidence Accurate, valid evidence Reliable evidence Evidence from the interpretation of their collected observation/data	Strong connection Evidence supports claims strongly Students provide evidence for most of the generated claims Evidence is strongly supporting claims even though it is about some claims Evidence is focusing on all the claims and strongly connected with claims	Strong understanding of how the evidence relates to the claim Student understands how their investigations tie into concepts about what they have learned in class Student make some connections to concept and real life Student spot errors and has some explanation for them
5	Multiple questions which include more than one open-ended questions Testable/scientific questions Questions catch the essence of the investigation thoroughly Questions are very significant and	Multiple claims Claims are from and based on the interpretation of their experimental observation/data (Claims about what they found out) Claims catch the essence of the investigation thoroughly	Very strong connection Questions and claims fit very strongly together Student develops claims for all the generated questions and all the provided claims are obvious in answering	Very powerful evidence Very accurate, valid, rich evidence Very credible and reliable evidence Evidence from interpretation of their collected observation/data	Very strong connection Evidence very strongly, effectively, and thoroughly supports the proposed claims Students provide evidence for all the generated claims	Thorough explanation how the evidence relates to the claim Student strongly understands how their investigations tie into concepts about what they have learned in class Students refers some real life application to make a connection with their laboratory work

	adequate Questions are of very high quality	Claims are very significant and adequate Claims are very valid, sound, and accurate Claims are of very high quality	questions		Evidence is very strongly supporting all the claims	Students have suggestions for correcting their errors Student recognizes what new things they have to think about

Modified from Choi, Notebaert, Diaz, and Hand (2010)

Scoring Matrix for the Holistic Argument

Points	Criteria
2	Very weak argument Untestable questions, invalid claims, and unreliable evidence Very weak connections between questions, claims, and evidence Do not flow smoothly from one area to another
4	Weak argument May be untestable questions, invalid claims, and unreliable evidence May not have reflection Weak connection between questions, claims, and evidence May not flow smoothly from one area to another
6	Moderate argument May be significant questions, adequate claims, appropriate evidence, and reflection Moderate connections between questions, claims, and evidence May flow smoothly from one area to another
8	Powerful/enriched argument Significant questions, valid claims, strong evidence, and meaningful reflection Strong connection between questions, claim, and evidence Flow nicely from one area to another
10	Very powerful/Enriched argument Essential questions, very sound claims, very strong evidence, and very meaningful reflection Very strong connection between questions, claims, and evidence Flow very nicely from one area to another

ACORN Student Sheet

ACORN

Acronym	Concept	Activity
A	Answer	**Answer** the question or address a given scientific claim.
C	Cite	**Cite** evidence related to the question or claim to support their answer. Evidence can come in the form of data from a lab, prior knowledge, or content knowledge.
O	Organize	**Organize** reasoning with citied evidence. Make a plan so you know your reasoning relates!
R	Respond	**Respond** to the question or claim using your evidence and reasoning!
N	Note	**Note** how you think you did in answering the claim or question.

From ACORN,
Science understanding grows…

ACORN Graphic Organizer #1

ACORN Organization Chart

Claim:

Evidence	**Reasoning**
Observations or examples that relate to the claim.	*Science ideas that explain how the evidences makes the claim true or false.*
Where did this come from? ❑ Lab Data ❑ Personal experience ❑ Example ❑ Chart ❑ Picture/Video ❑ Other: ______________________	Did I accomplish the following: ❑ My reasoning is based in science fact ❑ My reasoning makes use of the evidence ❑ My reasoning relates to the claim ❑ My reasoning is strong. ❑ My reasoning is detailed. ❑ My reasoning is clearly worded.

ACORN Graphic Organizer #2

ACORN Graphic Organizer

A *Answer the question by making a claim*

C *Cite evidence to support your claim*

1.

2.

3.

O *Organize your reasoning for each piece of evidence*

1.

2.

3.

R *Respond by using your evidence and reasoning to address the claim*

N *Note how you think you did by answering the following question:* ***Insert question appropriate for prompt here***