THE RELATIONSHIP BETWEEN SELF-ASSESSMENT AND CONSTRUCTING

EVIDENCE-BASED EXPLANATIONS IN A MIDDLE SCHOOL SCIENCE

CLASSROOM

by

Dalton Cox McCurdy

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

of

Masters of Science

in

Science Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2013

STATEMENT OF PERMISSION TO USE

In presenting this professional paper in partial fulfillment of the requirement for a master's degree at Montana State University, I agree that the MSSE Program shall make it available to borrowers under the rules of the program.

Dalton Cox McCurdy

July 2013

iii

TABLE OF CONTENTS

LIST OF TABLES

v

LIST OF FIGURES

ABSTRACT

This study investigated the link between formative assessments use and the ability to construct evidence-based arguments. Self-evaluation and other formative assessments were incorporated into daily lesson plans over a 4-week period to determine if these tools had a positive impact on writing outputs in 8th grade science. Prior to the intervention, a writing baseline was obtained by the use of an online writing prompt program as well as through teacher-graded rubric based assessments. The treatment was followed by a second round of benchmark testing using the WPP in congruence with teacher and student assessment rubrics. Data were also collected by means of student surveys and student interviews. All data were analyzed to determine whether there was a significant relationship between the use of student assessment tools and the ability to incorporate evidence-based explanations into formal and informal science writing assignments. Secondary research questions investigated if self-assessment use improved a students' self-awareness in regards to their content comprehension. Though results varied and a more thorough and extensive investigation would be necessary to pull out more specific trends, qualitative and quantitative data suggested that incorporating more self-assessments and formative assessments improved students' awareness in their writing outputs. It appeared that the more exposure students had to the process of persuasive writing, through exposure to writing samples and clear formulaic tools to assist in the process, the more comfortable students become in their abilities to construct these types of outputs.

INTRODUCTION AND BACKGROUND

As an eighth grade physical science teacher I have often noticed a disconnect between what students think they know and what they are able to get down on paper. Self-awareness still seems to be developing at this stage of formal education. Self-evaluation is a tool that I have always thought imperative to learning and one that I have hoped to incorporate more into my classroom.

Graland Country Day School is a private independent kindergarten-8th grade school located in Denver, CO. Caucasian students make up 79% of the Graland student body. Though Graland is a private school, 16% of students at the school are on financial aid. A majority of the eighth grade students have been at the school since kindergarten. All live within a 30-minute radius of the school.

My eighth grade physical science curriculum is lab-based with a large inquiry component. Throughout the first half of the year students investigate the fundamentals of chemistry, with a focus on classification of matter, physical and chemical changes, and atomic structure. Throughout the second half of the year students take on an introductory physics curriculum. Motion, forces, and simple machines make up the major focus. Throughout the entire year students are asked to complete an assortment of both formal and informal experiments and use collected data to justify and explain their findings. Most of this justification occurs in writing through formal and informal lab report write-ups as well as blog post reflections.

I have continuously noticed a struggle in these writing samples to back claims

clearly with evidence and link the two through persuasive reasoning. This observation was backed by recent test results on a national standardized test taken by all Graland eighth grade students. The ERB Writing Assessment Program (WrAP) assessment is a base-line writing assessment administered at Graland to all students in grades 2-8. In this assessment, students are given a prompt and asked to answer the question in essay form. Writing samples are then assessed and scored in 6 categories: overall development, organization/support, sentence structure, word choice, and mechanics. Students at Graland Country Day school scored lowest in the "support" category. These results reflect a struggle to support claims with clear and well thought out evidence.

In this study I wanted to see if an incorporation of self-evaluation and formative assessment tools into my daily curriculum would have a positive impact on support writing outputs. I specifically looked at the ability of students to construct sound explanations using evidence-based arguments and reasoning skills. While doing this, I also looked at secondary questions to determine if self-assessment use improved a students' self-awareness in regards to their content comprehension.

My primary focus question for this action research was:

1. Does the use of self-evaluation and formative assessment tools have a positive impact on support writing outputs in 8th grade Physical Science?

My secondary questions for this action research were:

1. Does student self-awareness improve with the use of self-assessments?

2. What type of self-assessments do students find the most useful?

CONCEPTUAL FRAMEWORK

Feedback plays a crucial role in the learning process (Andrade & Valtecheva, 2009). Research has shown that feedback promotes learning and achievement (Brinko, 1993; Butler & Winne, 1995). That being said, many students do not get an abundance of feedback simply because many teachers cannot respond to every piece of work completed by a student, or they can't do so timely enough to make a difference in the student's learning process. When students are able to evaluate their own work with a defined set of criteria or simple check-in questions, they are not only able to take learning into their own hands, but they are also able to monitor their own progress.

As a result of self-assessment, both self-regulation and academic achievement can increase (Schunk, 2003). When this tool is incorporated there is a clear shift in focus. Expectations are clear and students are provided with a clear set of guidelines to work towards. Students are seen having a proactive rather than reactive role in generating and using feedback in the learning process. This formatting allows for increased, and more personalized, regulation of the process. There is clear empirical evidence that shows that learners who are more self-regulated are more effective learners: they are more persistent, resourceful, confident, and higher achievers (Pintrich, 1995; Zimmerman and Schunk, 2001). Self-regulated learning can help personalize learning trajectories (Bjork, 1999). It allows students to better identify their strengths and weaknesses and then work toward improving upon them.

Formative assessment is used to improve the quality of student learning rather than just to assess students in order to evaluate their performance. (Angelo & Cross, 1993). Formative assessments are specifically intended to generate feedback on performance in order to improve and accelerate the learning process (Sadler, 1998). These evaluations can take many forms, but the ultimate goal of all these assessments is to cultivate a greater awareness of understanding. That being said, if formative assessment is completely in the hands of the teacher, it is hard for students to become empowered and take learning into their own hands outside the walls of the classroom (Nicol & MacFarlane-Dick, 2006). Formative assessment in the hands of the students is beneficial in molding more prepared learners. Especially when formative assessments are frequent and diagnostic in nature, they can help students develop meta-cognitive skills such as the ability to think holistically and to identify gaps in understanding (Steadman, 1998).

One type of formative assessment, rubrics, can teach as well as evaluate. According to Sadler (1989), there are three criteria that are necessary for helping students improve. First, expectations of work must be clear and articulated. Second, self-assessment must be in place. Third, after self-assessment there must be an opportunity for revision. These three steps alone have been associated with improvements in learning, particularly within the area of writing (Andrade & Valtcheva 2009)). Expectations of work can be made incredibly clear through the use of criteria-referenced self-assessments. These types of rubrics are well-organized, and easy for a student to follow. They can also serve as a clear guide for students in the revision process (Andrade & Valtcheva, 2009).

Past research has already linked criteria-referenced self-assessment and improvements in writing. A study of seventh and eighth grade students' writing by Andrade and Boulay (2003) found a positive relationship between self-assessment and quality of writing, especially for girls. An earlier study by Ross, Rolheiser, and Hogaboam-Gray (1999) described a scenario involving fourth, fifth, and six graders Students with weaker writing skills who were trained in self-assessment of narrative writing outperformed weak writers in the comparative group. The improved post-test scores of the group trained in self-assessment showed specific improvements in "integration of story elements around a central theme" (p. 124).

The purpose of my study was to continue to investigate this above-mentioned link between writing development and self-assessment tools. Evidence suggests that with improved self-awareness and more self-regulation as a learner, students are more successful in the classroom. My goal is to see if these "successes" are reflected in a student's writing. I was specifically interested in seeing if there was a connection between self-assessment and other formative assessment tools and a student's ability to create and construct evidence-based arguments in writing assignments.

METHODOLOGY

My goal with this study was to help students improve self-awareness in the classroom, specifically in the field of writing. The role of feedback is crucial in student learning. It provides students an opportunity to manage and self-regulate their own

learning. On a practical level, it also allows students to obtain more feedback than they

would from a teacher. As students become more in tune with what they truly

comprehend in the classroom, self-awareness and self-confidence will improve. As this

occurs they will be able to better express their thoughts and analysis through writing.

This connection was the main focus of my study.   I looked specifically at the students'

ability to use evidence in constructing explanations in writing. A secondary question

examined if student self-awareness improved with the incorporation of self-evaluations

and formative assessments.   Another secondary question investigated what type of self-

assessments students find most useful in this process.

The participants in this study were 59 eighth grade students at Graland Country

Day School: 34 girls and 25 boys. The eighth grade students completed a Physical

Science curriculum. The first half of the year focusesd on fundamentals of chemistry and

in the second half the year the focus shifts to fundamentals of physics. Daily lesson plans

were inquiry-based and demonstration-heavy. There was a large hands-on component to

the class as well. Students were assessed through nightly homework assignments, group

work, lab write-ups, formal assessments, projects, and daily participation. Students also

maintained e-portfolios set up on the blogger.com platform. Through these forums as

well as through lab write-ups of all kinds, writing was emphasized and practiced on

almost a daily basis.

This research study was incorporated into the 8th grade Physical Science

curriculum over a 4-week period during the 2013 winter semester. The topics covered

during this time period were motion and forces with a specific focus on friction, free-fall

and balanced vs. unbalanced forces. The real-world application of this content was emphasized throughout the unit.

Prior to the first week of the intervention, a lesson was incorporated into the curriculum on claim-based reasoning. The concept of "C.E.R" was introduced which helps breakdown reasoning into a Claim + Evidence + Reasoning =Explanation (E. Brunsell, Personal Communication, September 23, 2012). This idea was practiced as a group. We broke down effective writing into the above-mentioned categories and also identified C.E.R. material necessary to back a claim (ex: *how do you know the unidentified substance is a mixture?*). Signage for C.E.R was also placed on a large poster in the front of the classroom (CLAIM +EVIDENCE+REASONING = EXPLANATION) and smaller signage was given out to each student as a reference. Self-evaluation and formative assessment techniques were also discussed and explained in order to make sure that expectations were consistent throughout the class.

Over the four-week time span a multiple step intervention was incorporated into the curriculum. During the treatment period, formative assessments were incorporated into the curriculum two times per week (Appendix B). The assessments provided an opportunity for a comprehension check-in. One-minute papers were included in the treatment in order to help students prioritize and synthesize what they learned.   These short write-ups were returned to students so they could reference them as necessary when writing up their formal lab reports.

C.E.R recognition techniques were also incorporated. Students were asked to write down the main points of each lecture and identify evidence used to reinforce each

claim. Over the second half of the treatment period, students worked independently on a project looking at the various forces that impact a ski jumper. Daily references to C.E.R were made and students reflected on their daily work on their blogs. Students were also encouraged to look at each others blogs at night on order to make sure their own work was through, explanation-based, and up to class expectations.

A criteria-referenced self-assessment rubric was incorporated into the project write-up as well. Students also completed a data and analysis write-up on their project work on their blog. When using this evaluation, students compared their work to explicitly stated criteria, goals, or standards (Appendix C). They were asked to identify if writing samples appropriately used evidence and reasoning to back their claims. After these evaluations were completed, students were given the opportunity to revise their work accordingly prior to formally handing their work in.

Data were collected through a variety of means before and after the intervention. Prior to initiating treatment, the ERB Writing Program: Writing Practice Program (WPP) a computer assessment tool, was used to help obtain a pre-treatment baseline for writing outputs. This assessment tool scoring techniques correlate with the ERB WrAP Assessment that students take formally once a year in grades 2-8. Within this testing program, I was able to provide a prompt for students to write about (Table 1). Students were given a 30-minute time cap to create their output. Use of "support" and overall organization were ranked on a scale of 1-6, with one being the lowest and six being the highest (see Appendix A). Scores are calculated instantaneously and students are able to view their outputs right after they click "Submit". After the 4-week treatment, a post-

treatment assessment was also completed using WPP in order to provide a comparison.
Within these computer-based assessments, "use of support" scores were compared.
Prompts for the pre and post-treatments were both based on similar length inquiry-based
lab exercises completed in class. On both occasions students were asked to write a
formal lab report. The lab report could be used during the prompt write-up in order to
reference and incorporate specific data.

Before and after the treatment period, students' confidence in their own writing
skills was also surveyed. An online survey (Appendix D) was given to the students prior
to the intervention in order to obtain baseline opinions of self-evaluation use. Student
surveys were used once again after the treatment (Appendix E). The same questions
from the pre-treatment survey were asked again in order to obtain comparative data.
Questions about specific formative assessments used were also incorporated into the
post-treatment survey.

Table 1
WPP Prompts

Pre-Treatment Prompt	You just completed a lab investigating how fast you can scooter on two different surfaces. What factors affected your speed? Please be specific in your response.
Post-Treatment Prompt	You just completed an activity creating the most "successful" ski jumper. What factors contributed to the success of your prototype? Please be specific in your response.

In addition, two students were selected at random from each class (eight in total) to be
interviewed to obtain their thoughts on the treatment as well to gather additional student

opinion-based evidence on use of formative assessments and their effectiveness in

improving writing skills (Appendix F).The data collection techniques are summarized in

Table 2.

Table 2
Data Triangulation Matrix

Focus Questions	Data Source 1	Data Source 2	Data Source 3
Primary Question: 1. Does the use of self-evaluation and formative assessment tools have a positive impact on support writing outputs in 8th grade Physical Science?	Pre and Post Treatment WPP Computer Assessment outputs	Student surveys	Student interviews
Secondary Questions:			
2. Does student self-awareness improve with the use of self-assessments?	Student surveys	Student interviews	Informal class Observations
3. What type of self-assessments do students find most useful?	Student surveys	Student interviews	E-portfolio reflections and classroom observations

The research methodology for this project received an exemption by Montana State

University's Institutional Review Board and compliance for working with human subjects

was maintained.

DATA AND ANALYSIS

Data were collected throughout the study as planned. Writing samples were assessed by means of the ERB Writing Program: WPP pre and post treatment. Student surveys as well as post-treatment interviews were also used to analyze treatment effectiveness.

Data from the ERB Writing Program: Writing Practice Program (WPP) Assessments did not show a significant difference between pre and post treatments outputs (Figure 1). Students scored an average of 18.94 prior to the 4-week treatment period and a 19.36 after the treatment phase was completed.   That being said, the standard deviation for treatment data was 3.08. The standard deviation for post-treatment data was even greater (4.31). These high deviations suggest that assessment outputs varied quite a bit and the treatment did not have a consistent effect on all those involved.
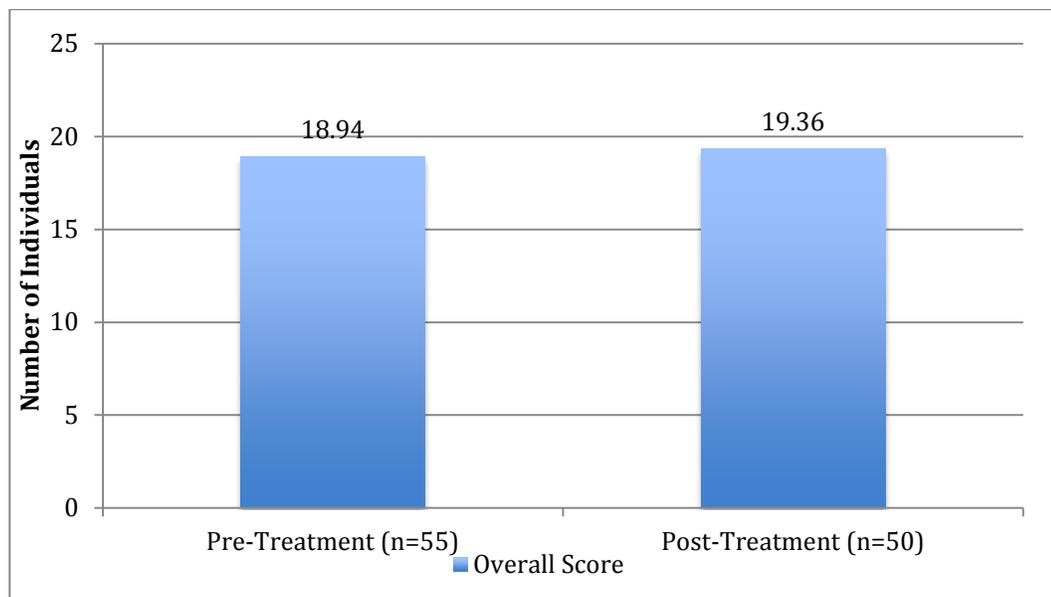

Figure 1: ERB Writing Program: Writing Practice Program (WPP) assessment outputs.

Looking at a breakdown of the WPP output results, again no significant trends appeared in the data (Table 3). Support outputs in particular were quite consistent (+.05 difference between results). Again, however, standard deviations were quite high. The pre-treatment data had a standard deviation of 0.62 and the post treatment data had a standard deviation of 0.77.   These large variances among the data sets once again suggest that the treatment had varying impacts on the participants. It should be noted, however that, after a closer look at WPP scores, there were 9 student scores that had significantly lower scores on the second WPP assessment. These 9 students had an average decline in score of 5.3. This second WPP writing assessment was assigned as a nightly assignment (due to time restraints in class) while the first WPP writing assessment was completed in class. The timing of the second assessment was also less than ideal as it occurred right before the close of a marking period and students had quite a bit on their plates at this time. This extra-stress and heavy workload may have affected some students' inputs on this second writing assignment.

Table 3.
ERB Writing Program: Writing Practice Program (WPP) Assessment Outputs Broken Down by Category.

	Pre-Treatment (n=55)	Post-Treatment (n=50)
Overall Development	3.11	3.16
Organization	3.21	3.14
Support	2.79	2.84
Sentence Structure	3.18	3.48
Word Choice	3.21	3.24
Mechanics	3.44	3.5

It should also be noted that throughout both testing periods, a bias was noted with WPP results. Outputs scores seemed to be higher with increased length of written work put into the program. Longer write-ups seemed to produce higher scores. This makes sense, as this software is most commonly used to grade formal 5-paragraph essays. Students in the study are quite familiar with this program and use it in English class quite often. All students noted a significant drop in their output scores. That being said, the decrease seemed to be consistent across the board and should not have an effect on comparison of pre and post treatment data.

In order to put to the WPP results in a larger context as well as get another measurement on student learning to see if improvement were indeed made in writing over the course of this study, I compared pre and post samples of writing of Student X who improved by 4 points on the WPP assessment after the treatment period (Table 4) as well as Student Z whose WPP results decreased by 2 points after the treatment (Table 5). Looking strictly for examples of C.E.R, both samples appeared to improve in output. In post treatment samples of both students, reasoning seemed to be improved and more methodical in format. Claims were followed more consistently by evidence and reasoning connecting the claims made to concepts. This suggests that, though WPP results did not reflect an improvement in support writing, if we look specifically at C.E.R use and define an increase in use of that model as a positive result, the treatment period may have had a more positive impact than that one learning measurement reflects.

Table 4.
Pre and Post Treatment Writing Samples for Student X

Pre-Treatment Sample	Post Treatment Sample
Instead of separating it physically with distillation, filtration, or evaporation, we do a chemical change. First, we started with a beaker full of water. We then put a test tube of water in each beaker and flipped it over but assured that no water escaped. There were two probes inside each test tube and they were connected with cables, which were attached to an ordinary electrical outlet. To create a more conductive water, we added Epson salt. Because of this, bubbles started occurring in both test tubes. Gas was then forming at the top of each tube. It was noticed that one gas bubble was twice as large as the other. The ratio for water is 2:1.	*It appeared that my prototype had clear lift off the ramp. On the car, I lifted the foam board on one side so one was higher than another. Then I put the higher side facing down on the ramp, so when it flew off the air would have air resistance. Because of this, the car could get height in the air. Also, because of the aero dynamics of the plane and the wheels on the car, there was enough speed to get off the ramp due to the minimal amount of friction (rolling and fluid). Adding to this was the tilted foam board, which would give it lift.*

Table 5.
Pre and Post Treatment Writing Samples for Student Z

Pre-Treatment Sample	Post Treatment Sample
We appeared to collect more of one gas than the other because the two gases were different. One of the gases was hydrogen, and the other was oxygen. The test tube with more gas was hydrogen, while the one with less gas was oxygen.	*Our balloon car prioritizes speed because the rubber band car is faster than the balloon car. We chose a lightweight design, and a big rubber band for propulsion. We want minimize drag. We will maximize air resistance by making the sides of the car angled out to the wheels. Newton's second law states that acceleration equals net force over mass. We will maximize acceleration by creating a small mass (a lightweight design). To propel the car forward, we will create a slingshot with our hands and use the Lego to hold it.*

Results from pre and post treatment student surveys also suggested that students' opinions on writing and self-awareness shifted throughout this study. Student Responses to the Statement *I am confident in my ability to connect specific evidence to reasoning*

and claims made in science class showed a clear shift in confidence (Figure 2). Students

that "completely agreed" with this statement increased from 11% to 26% after the

completion of this study. Students who "strongly disagreed" with this statement

decreased from 5% to 2%. This is a clear indicator that, though WPP outputs may not

have reflected an improvement in students' ability to connect specific evidence to

reasoning and claims after the treatment period, student awareness and confidence in this

realm improved.
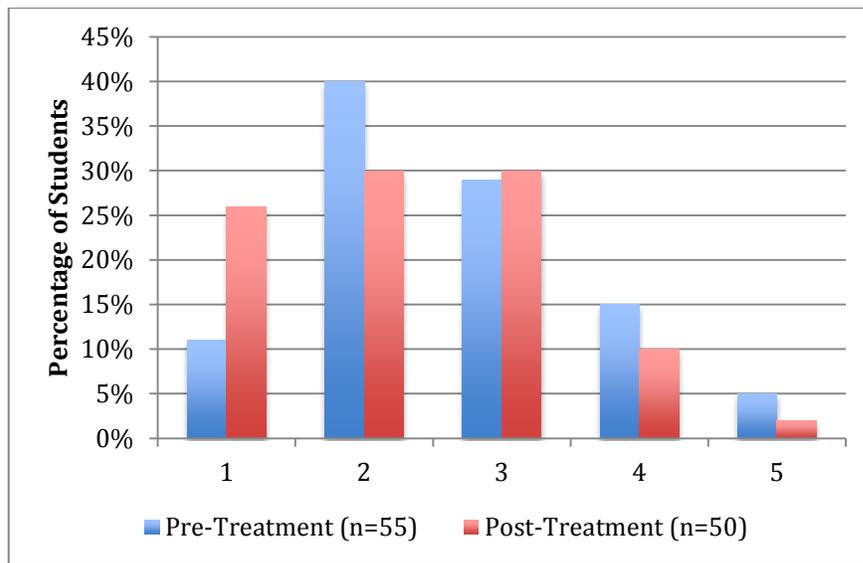
Figure 2. Student responses to the statement *I am confident in my ability to connect specific evidence to reasoning and claims made in science class (1 = Strongly Agree, 5 = Strongly Disagree).*

Student Responses to the Statement *I am successful at constructing and utilizing*

evidence-based claims in my writing reaffirm the findings of above (Figure 3). The data

again show a shift toward more agreement with this statement. The percentage of

students prior to the treatment period who "completely agreed" with this statement was

9%. Post treatment survey results show that this percentage increased to 22%. Students

who "completely disagreed with this statement declined from 4% to 2%. As with the question addressed above, this data suggest an increase in students' confidence levels in regards to their writing capabilities.
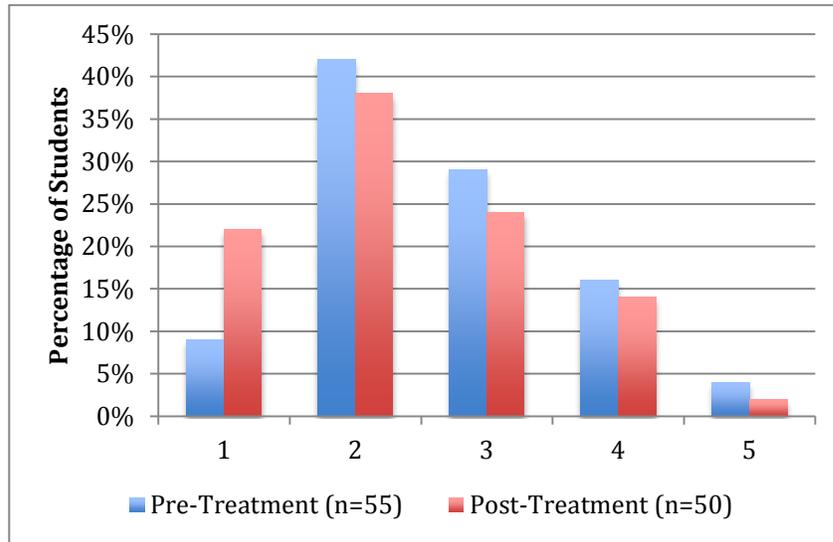
Figure 3. Student responses to the statement *I am successful at constructing and utilizing evidence-based claims in my writing (1 = Strongly Agree, 5 = Strongly Disagree).*

Students' opinions on the usefulness of self-assessments also shifted during this study. Student Responses to the Statement *Self-assessments are helpful in improving my understanding of content covered in science class* became more negative after the treatment period was completed (Figure 4). The percentage of students who "somewhat agreed" with the above statement decreased from 33% to 10%. The percentage of students "neutral" to the above statement increased from 24% to 38%. The percentage of students to "slightly disagree" with the above statement increased from 15% to 24%. This data also correlate with general class observations taken during the times when self-evaluation forms were handed out and completed. Students were visibly and vocally uncomfortable evaluating their own work. One student admitted to having a difficult

time with self-evaluation because "it's hard and embarrassing to critique your own work".

Others agreed with this statement. Because of this discomfort with self-evaluation, many

appeared to not take these evaluations very seriously and rush through them. Especially

with the rubric work, many students simply circled the boxes in the rubric that they

thought applied to them and did not take the time to take down any notes or comments.

As a result of this, many students completed the process in a very short time period.
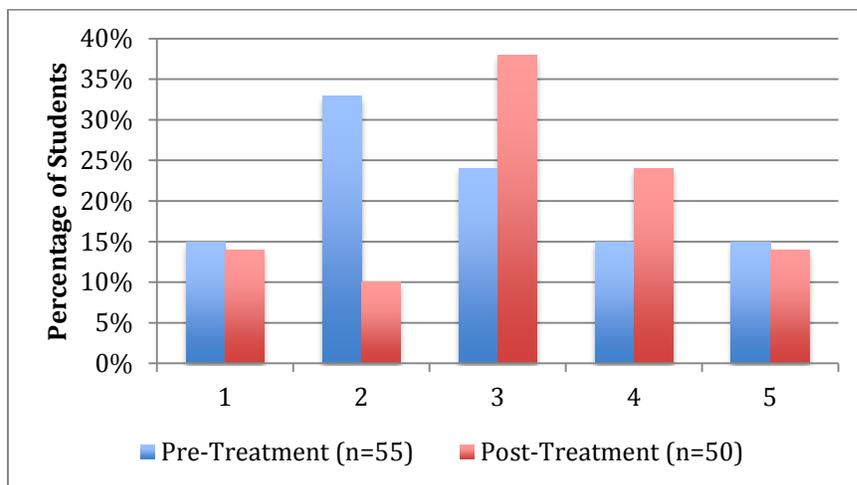
Figure 4. Student responses to the statement *Self-assessments are helpful in improving my understanding of content covered in science class (1 = Strongly Agree, 5 = Strongly Disagree).*

Student Responses to the Statement *Self-assessments are helpful in improving my*

writing skills in science were fairly inconclusive and no significant trends were seen in

the data (Figure 5). Though there was a slight shift in the percentage of students who

were in complete agreement with the above statement (an increase from 9% to 14%) the

majority of responses remained fairly neutral.
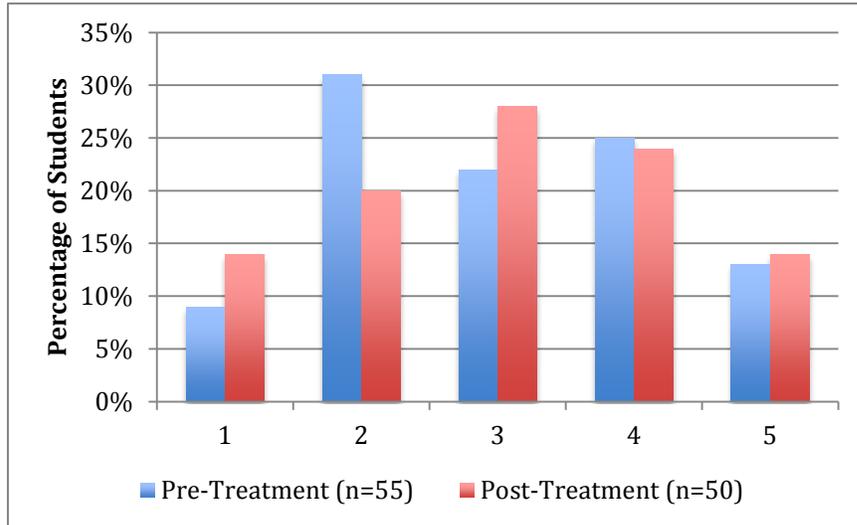
Figure 5. Student responses to the statement *Self-assessments are helpful in improving my writing skills in science (1 = Strongly Agree, 5 = Strongly Disagree).*

Student Responses to the Statement *I am confident in my ability to connect specific evidence to reasoning and claims made in science class* showed a slight bimodal shift after the treatment period (Figure 6). In pre-treatment responses most students appeared to "somewhat agree" with this statement or be "neutral" toward it (65% of participants). After the treatment, the percentage of students who completely agreed with this statement increased from 11% to 14%. The percentage of students who completely disagreed with this statement increased from 7% to 16%. The number of students that "somewhat agreed" or remained "neutral" on this question after the treatment phase declined to 50% (15% decrease). This suggests that students became more opinionated on this question (either positively or negatively) after the treatment. Though the treatment did not necessarily help all students become more comfortable evaluating their work, it definitely raised awareness of individual students' abilities in this area.
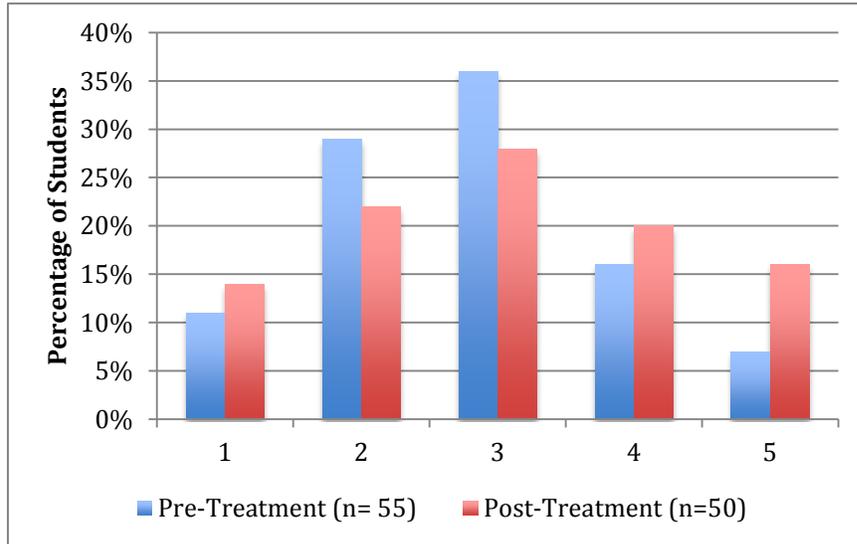
Figure 6. Student responses to the statement *I am comfortable and confident evaluating my own work by means of a self-assessment (1 = Strongly Agree, 5 = Strongly Disagree).*

Student Responses to the Statement *I have a good idea of what I know and what I need to work on in science class* also remained fairly consistent before and after the treatment period (Figure 7). The only slight shift appeared to be with those who initially "somewhat disagreed" with this statement. Percentages in this category decreased from 13% to 6%. This suggests that self-assessments might have been a more useful tool for those who initially were struggling evaluating their own work. Those who initially were comfortable and confident with this approach remained so.
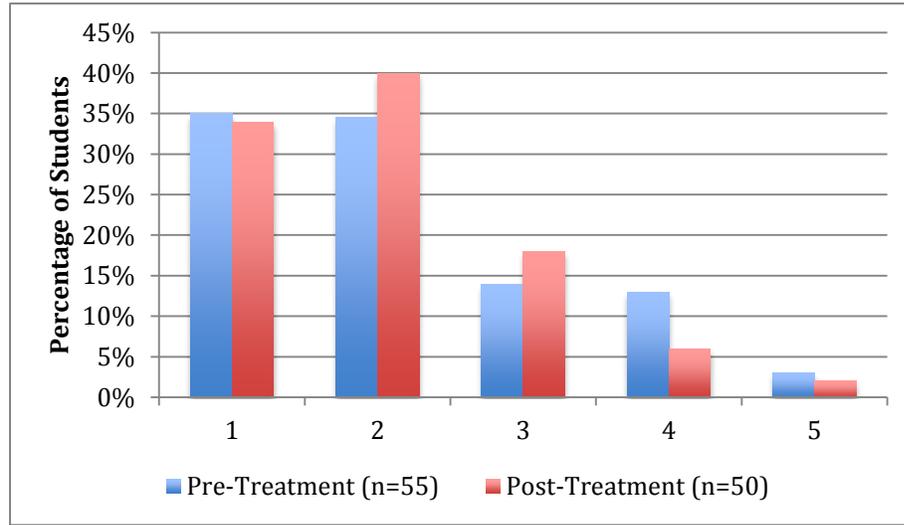
Figure 7. Student responses to the statement *I have a good idea of what I know and what I need to work on in science class (1 = Strongly Agree, 5 = Strongly Disagree).*

Post-treatment surveys suggest that half of the students were fairly neutral toward the use of the one-minute essay as effective formative assessment tool in helping synthesize content and learning (Figure 8). Class observations appeared to correlate with these data. There was not a clear vocal or visual response to these short in-class assignments.
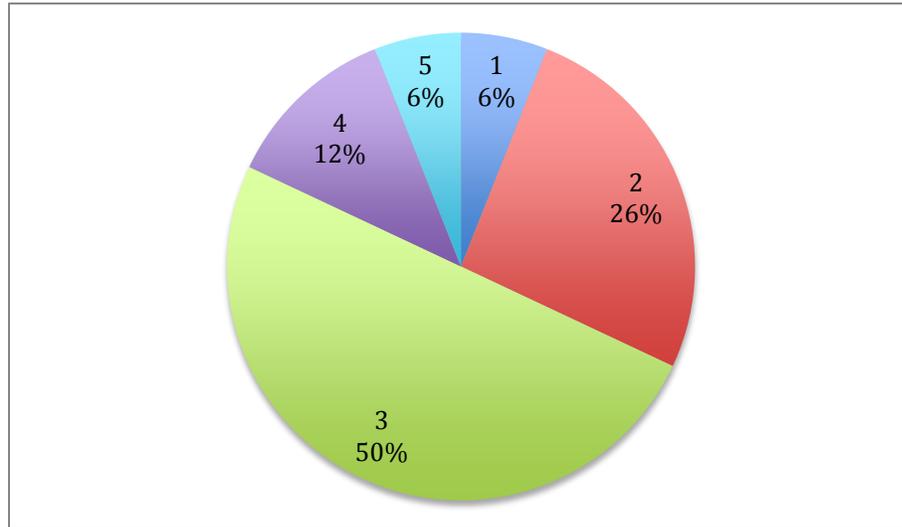
Figure 8. Student responses to the statement *the one-minute essay helped me better synthesize material I've learned in science (1 = Strongly Agree, 5 = Strongly Disagree).*

Post-treatment surveys suggest that the majority of students (74%) were fairly neutral or in slight disagreement to the statement *the one-minute essay helped me better connect evidence with claims in science class* (Figure 9)
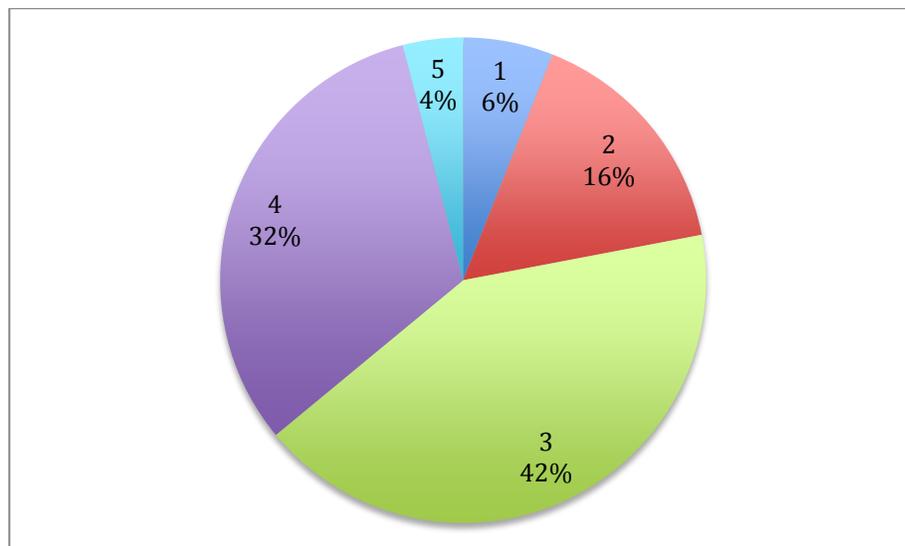
Figure 9. Student responses to the statement *The one-minute essay helped me better connect evidence with claims in science class (1 = Strongly Agree, 5 = Strongly Disagree).*

Student Responses to the Statement *using self-evaluation forms helped me identify some of my weaknesses in writing* also appeared to be fairly neutral in nature (Figure 10). 70% of participants were "neutral", in "somewhat agreement" or in "somewhat disagreement" with this statement. 14% of students strongly agreed that self-assessments could be helpful in identifying personal weaknesses in writing. A slightly higher percentage (16%) strongly disagreed with this statement. This incredibly varied data set suggests that the effectiveness of self-evaluations is quite individualized. It should also be noted that the degree to which these evaluations were utilized also varied quite a bit. Many students looked at the evaluations through a very critical lens. Others completed the process incredibly quickly. This tool was given out as an aid to improve student writing and thus was not collected or graded. As a result of this, many students took this assignment more seriously than others. These variations in approach could have had an impact on outputs and opinions. If the treatment period was expanded and self-evaluation became a more regular part of the routine, results might have proved more significant.
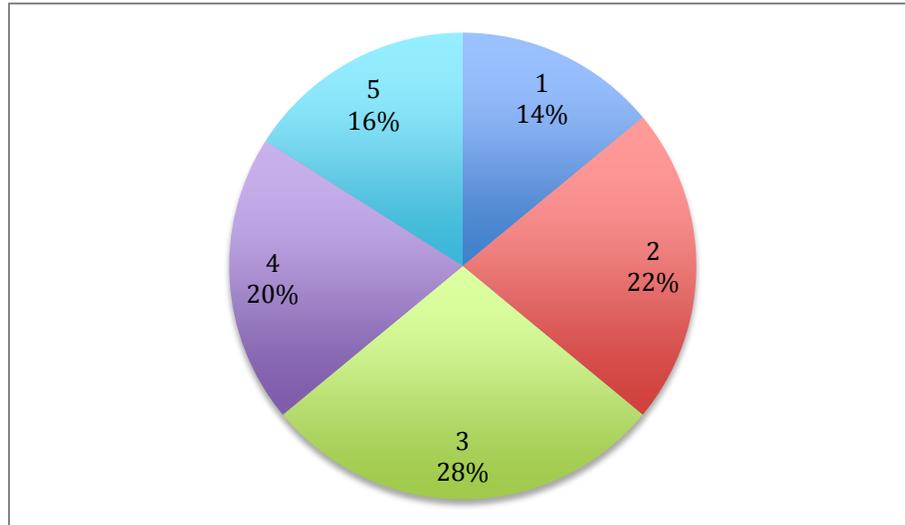
Figure 10. Student responses to the statement *Using self-evaluation forms helped me identify some of my weaknesses in writing (1 = Strongly Agree, 5 = Strongly Disagree).*

The majority of student responses to the statement *self-evaluation has improved my comprehension of content covered in this class* were neutral or in slight agreement (Figure 11). 58% responded with a 2 or 3 on the outlined scale. This data set was fairly similar to responses to the statement *using self-evaluation forms helped me identify some of my weaknesses in writing.* There appeared to be a few more people who were in slight agreement with this statement as opposed to the latter. Informal class observations do suggest that students became more comfortable and more astute at identifying weaknesses in their writing, particularly when it came to connecting clear evidence to claims. I heard many people using the term "C.E.R." when discussing their work with me as well as with their peers during independent work time.
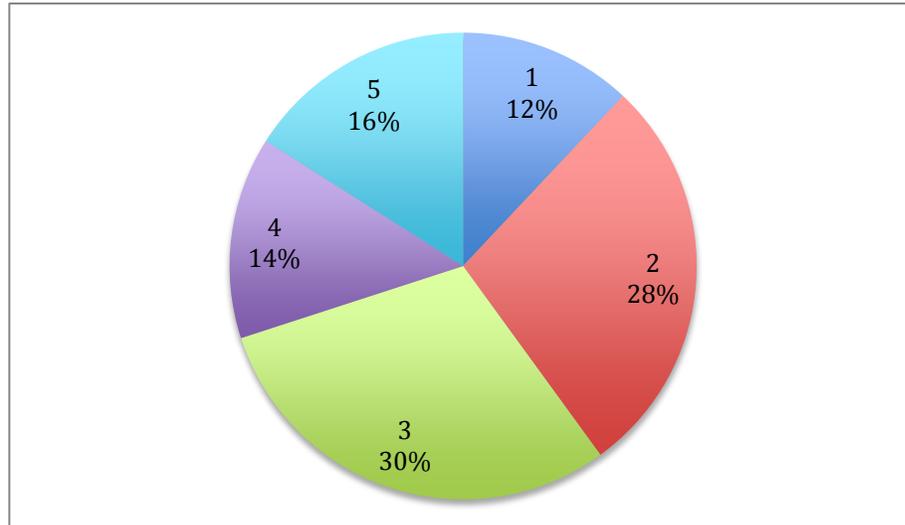
Figure 11. Student responses to the statement *Self-evaluation has improved my comprehension of content covered in this class (1 = Strongly Agree, 5 = Strongly Disagree).*

Responses to the statement *after utilizing the self-assessment rubrics, I was able to improve my writing outputs in this class* also seemed to be quite varied (Figure 12). Though the majority of students seemed to agree with this statement, answers were again fairly evenly distributed across the scale. It would be interesting to see how these answers shifted if the treatment period was expanded and rubric use became a more routine form of self-evaluation. As noted above, often times it seemed that students took advantage of the criteria listed on rubrics and simply circled the area of best-fit without having to record or comment on the process. Because of this, it appeared that many rushed through this practice and did not put enough time or thought into the practice for it to be fully effective.
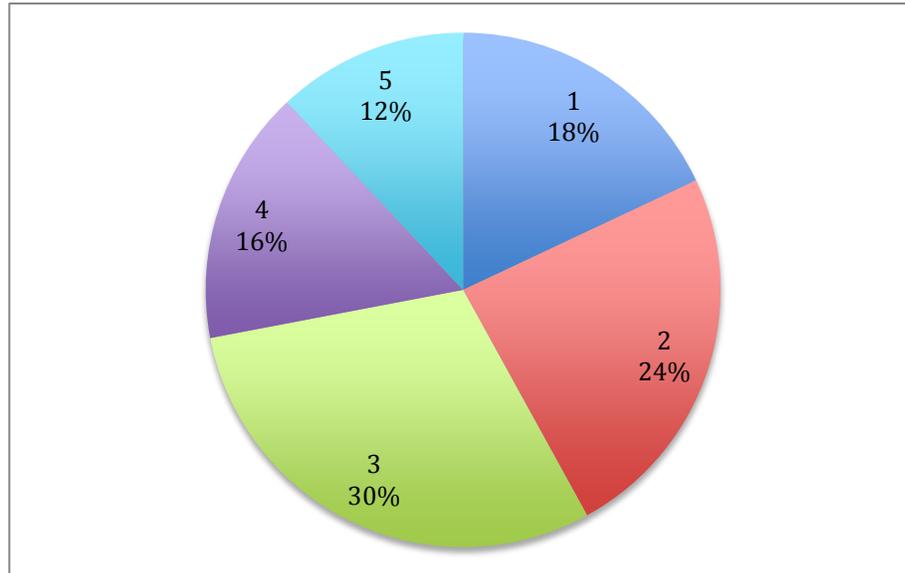
Figure 12. Student responses to the statement *After utilizing the self-assessment rubrics, I feel like I was able to improve my writing outputs in this class (1 = Strongly Agree, 5 = Strongly Disagree).*

Informal interview answers reiterated many observations made in the classroom setting throughout the treatment period. Students agreed that backing up claims with clear evidence was important in scientific writing. They agreed that evidence makes arguments more persuasive. That being said, they were apprehensive about self-evaluation work because they were uncomfortable with how their self-critiques would be perceived by their peers. Student A said, "I don't like using self-evaluation tools because I never want to look like I am grading myself to easy or too hard. I don't want others to think I'm full of myself". Student B said something similar, "self-evaluation tools make me uncomfortable. I really don't like grading my own work". Other students who responded had similar answers with the exception of one. Student C said, "I like self-evaluation tools. They work as a guideline to help me make sure that I've done my best job on science report write-ups". Though much of the discussion about self-evaluation

tools was negative, the majority of the students interviewed very much liked the use or C.E.R in putting together persuasive explanations. They found this approach formulaic, clear, and easy to use. Student D said, "I think using the idea of C.E.R really helped me better organize my writing in science". Student A noted, "Using C.E.R helps me make sure to include all important pieces to an argument or explanation. The consensus was clear that this approach simplified, and thus made students, more comfortable with this process.

In summary, student survey data collected pre and post treatment period appeared to be more conclusive than data collected only after the treatment period on the effectiveness of specific formative assessment tools. Results suggested that the more exposure students got to the process of persuasive writing, through clear vocal and visual reminders such as signage and daily referencing to the process, the more comfortable and confident they were with their abilities to connect claims with evidence. Though formal writing results did not show a significant shift in this area, students' awareness and confidence in their writing skills clearly improved. Student interview responses reiterated that students perceive persuasive science writing as being important and the use of clear evidence imperative. Students however, were very apprehensive to make full use of self-evaluation techniques due to peer perception of how individuals viewed themselves.

INTERPRETATION AND CONCLUSION

Overall it appeared that the more exposure students have to the process of

persuasive writing, through exposure to writing samples and clear formulaic tools to

assist in the process, the more comfortable students become in their abilities to construct

these types of outputs. It is also clear that this type of shift in writing, towards more

consistently including claims, reasoning, and evidence within explanations, does not

occur over night. It is a process, and one that students do not necessarily feel comfortable

immediately jumping into. Students are apprehensive towards completely utilizing new

formative assessment tools, especially when those tools involve self-assessment.

Critiquing their own work is not something that students are comfortable with. This may

simply be a function of lack of exposure.

The Impact of Self-Evaluation and Formative Assessment Tools on Support Writing

While the data collected in this study begin to look at the impact of self-

evaluation and formative assessment on writing outputs, a more extensive study, where

data collected over an extended period of time would be needed to really find stronger

correlations between formative assessment tools and support writing outputs.   An

extended study would also help eliminate some of the variations in data outputs that may

be a result of student stress and workload levels at various points throughout the year. In

a more extended study, I would also recommend that more varied learning measurement

tools be incorporated. Based solely on the data that was collected over the 4-week

treatment, the primary focus question is left unanswered. Though there are data that

suggest exposure to formative tools and self-evaluations forms improves student-

confidence, data on the effectiveness of the tools themselves are quite inconclusive.

What this suggests is that students are still apprehensive to use certain formative

assessment tools, simply raising awareness through means of daily discussion and simple

exposure to these tools does appear to improve student self-confidence.

Self-Assessments and Self-Awareness

It is difficult to connect a heightened self-awareness specifically with the use of

self-assessments. It is clear that with a larger focus on self-evaluation and continuous

feedback via peers or a teacher, students do seem to improve their ability to be self-aware

and more confident assessing their comprehension level. That being said, using formal

self-evaluation tools does appear to make students very uncomfortable. Many students

believed reflecting formally on their own work was too difficult. I'm not sure if it has to

do with where middle school students are socially or developmentally, but there was

definitely a fear with evaluating yourself too harshly *or* too softly in front of your peers.

This was definitely a concern. As a result of this, many students took an incredibly

neutral approach when using these tools.

Effectiveness of Self-Assessments

Because students are apprehensive towards assessing their own work in front of their peers, I am not sure if formal self-evaluations are the most effective tool to use at the middle school level. That being said, students do appear to be more comfortable comparing their work to others. If I were to take this Action Research study a step further, I would investigate the effectiveness of peer review as well as the effectiveness of comparing one's own work to their peers. We started doing this informally through student digital portfolios. After many students completed self-evaluation forms I had them look through the work of their peers electronically. This is also something that can be done outside the classroom. This could lessen the impact of social pressures. This type of comparison is a bit more concrete, but very clear. Students are clearly influenced by their peers; this type of review might be even more effective than a clear criteria-based rubric because it is based on real examples and is not so much of a formal self-review.

VALUE

This study has showed me the importance of formative assessment on student learning. Especially when it comes to truly understanding material and being aware of difficult content and or perhaps misconceptions, formative assessment work seems to raise student awareness and provide a forum for reflection on self-work. Students, especially at the middle school level, seem to benefit quite a bit from feedback. It also

seems that the faster than can obtain this feedback the easier it will be for them to apply it to their work. The WPP writing program used in this study provided students with this type of instantaneous feedback, and though there did appear to be a bit of a bias due to length of writing pieces, this feedback (or even the awareness that feedback was coming quickly) kept students thinking and on-task.

I was also a bit surprised by the pushback from students I received in the classroom regarding self-evaluation work. Students were much more comfortable having their teacher or other peers grade their work than when they had to self-evaluate. This definitely was not the response I expected. I think there is incredible value in being able to self-reflect and evaluate one's own work with a thoughtful and critical eye. As this type of tool becomes more common in the classroom, I think students will become more comfortable with the process. I also wonder if it would help to incorporate some type of incentive to keep students interested and motivated in using self-assessments. In the future, I was thinking about awarding extra credit to students whose self-evaluations were most similar to my own evaluations. This type of approach, almost making it a game, may shift the focus a bit and distract students from current fears concerning peers perceptions of how an individual student is grading him or herself.

Student perception of personal comprehension is definitely something that needs to be improved at the middle school level. If students can improve their understanding of themselves they will be able to better self-advocate as well as better personalize the reviewing process prior to summative assessments. As this is my first year at Graland Country Day School, I have definitely noticed this need to better self-evaluate in the

learning process. This has become more apparent to me as the year has gone on. That

being said, I probably would have structured this study a bit differently if I had seen this

trend a bit earlier. This study was an excellent step in identifying weaknesses and

strengths in the population I work with. I look forward to taking this new knowledge and

improving my own teaching strategies to better address these issues in the years ahead.

REFERENCES CITED

Angelo, T. A & Cross, K.P. (1993). *Classroom Assessment Techniques: A handbook for college teachers.* San Francisco, CA: Jossey-Bass.

Andrade, H., & Boulay, B. (2003). Gender and the role of rubric-referenced self-assessment in learning to write. *Journal of Educational Research, 97*, 21-34.

Andrade, H & Valtecheva, A. (2009) Promoting learning and achievement through self assessment. *Theory into Practice, 48*, 12-19.

Andrade, H., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effects of a model, criteria generation, and rubrice-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practices, 27*(2), 2-13.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher and A. Koriat (Eds.), *Attention and performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 435-459). Cambridge, MA: MITPress.

Brinko, L.T (1993) The practice of giving feedback to improve teaching. *Journal of Higher Education, 64,* 574-593.

Butler, D. & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65,* 245-281.

Nicol, D.J & MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback. *Studies of Higher Education Vol. 31*(21)*,* 199-218.

Pintrich, P. (1995). The role of goal orientation in self-regulated learning. In M. Boekarts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). San Diego, CA: Academic Press.

Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A. (1999). Effects of self-evaluation training on narrative writing. *Assessing Writing, 6,* 107-132.

Sadler, D. R. (1998). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119-144.

Schunk, D. (2003). Self-efficacy for reading and writing: Influence of meodeling, goal-setting, and self-evaluation. *Reading & Writing Quarterly, 19*, 159-172.

Steadman, M. (1998). Using classroom assessment to change both learning and teaching. *New Directions for Teaching and Learning, 75*, 23-35.

Zimmerman, B. & Schunk, D. (2004). Self-regulating intellectual processes and outcomes: A social cognitive perspective. In D. Dai & R. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 323-349). Mahwah, NJ: Lawrence Erlbaum Associates.

APPENDICES

APPENDIX A

ERB WRITING PRACTICE PROGRAM (WPP) SCORING GUIDE

ERB | WrAP WRITING ASSESSMENT PROGRAM

RUBRIC GRADES 3–12
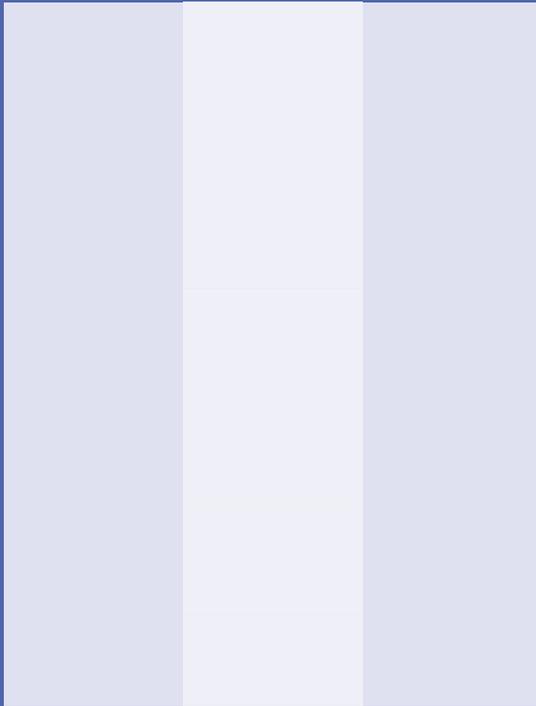
or complexity of
the piece (errors cause major
problems for readers)

- Mechanical errors that seriously
  interfere with communication
- Too brief to evaluate

©2010

*Grades 3/4 Narrative only

- NON-SCORABLE (N): is illegible: i.e., includes so many indecipherable words that no sense can be made of the response
  —or— is incoherent: i.e., words are legible but syntax is so garbled that response makes no sense —or— is a blank paper

APPENDIX B

FORMATIVE ASSESSMENTS USED DURING INTERVENTION

1. One Minute Paper Prompts

Prompt 1: take 1 minute to write down the three forces that affect an athlete ski jumping?

Prompt 2: take 1 minute to identify the main factors you will take into consideration when creating your ski-jumper prototype

Prompt 3: identify 2 pieces of evidence used to back claims made by your teacher in today's lecture

2. . C.E.R Prompts

On an index card identify a C.E.R used in class today
- claim:
- explanation:
- reasoning:

Using, C.E.R., make one explanation as to why your ski jumper didn't go farther on trial #1

3. Comprehension check-ins

Prompt: On a scale of 1-5 (5 being highest and 1 being lowest) How successful were you today at identifying evidence to support claims?

APPENDIX C

SELF-ASSESSMENT WRITING RUBRIC

Science Writing Rubric   Name:_____

	1	2	3	4	SCORE
Understanding of Content	Write-up shows a very limited understanding of (or serious misconceptions about) the topic. Relationships among science facts and concepts are not explained and no attempt is made to incorporate appropriate scientific language.	Write-up shows a partial understanding of the topics. There is an attempt to incorporate appropriate scientific terminology and explain relationships between science facts and concepts.	Solid understanding of the topics addressed in the assignment. Scientific terminology incorporated. Relationship among science facts and are explained.	In-depth understanding of the topics addressed in assignment. Appropriate scientifically terminology incorporated consistently throughout. Relationship among science facts and concepts are thou roughly and clearly explained.	
Use of Evidence	Minimal to no attempt made to incorporate evidence into arguments made in write-up.	A clear attempt was made to use evidence to back some of the claims made in the write-up.	Solid evidence used to successfully back a majority of the claims incorporated into the write-up	Very clear and well-explained evidence used to back claims incorporated into write-up	
Organization	Ideas are not presented in a clear or logical manner. All aspects of the assignment are not completely or satisfactorily addressed.	Ideas are not always expressed in a clear and logical manner, making the response difficult to follow. Various aspects of the assignment were only partially addressed.	Ideas are expressed, for the most part, clearly and succinctly. The various aspects of the assignment are generally addressed.	Ideas are expressed clearly and succinctly in a logical manner. All aspects of assignment are addressed.	
Spelling and Grammar	Major flaws in spelling and language conventions make the assignment difficult to follow.	Flaws in spelling and language interfere with the write-up.	Spelling and language conventions are generally correct.	Spelling and language conventions are correctly applied.	

Class:_____

Task/Assignment: _____

APPENDIX D

PRE-TREATMENT STUDENT SURVEY

PRE INTERVENTION SELF-ASSESSMENT SURVEY

Please answer the below survey questions honestly. Your participation or non-participation will not affect your grade or class standing. These surveys are anonymous. Participation is voluntary, and you can choose to not answer any question that you do not want to answer, and you can stop at anytime.

INSTRUCTIONS: Please respond to the statements below using the following scale 1-5 scale. **5-strongly agree, 4-Somewhat Agree, 3-Neutral, 2-Somewhat Disagree, 1-Strongly Disagree.**

1. I am confident in my ability to connect specific evidence to reasoning and claims made in science class.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

2. I am successful at constructing and utilizing evidence-based claims in my

    writing.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

3. Self-Assessments are helpful in improving my understanding of content covered in science.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

4. Self-Assessments are helpful in improving my writing skills in science.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

5. I am comfortable and confident evaluating my own work by means of a self-assessment.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

6. I have a good idea of what I know and what I need to work on in science class.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

APPENDIX E:

POST-TREATMENT STUDENT SURVEY

POST INTERVENTION SELF-ASSESSMENT SURVEY

Please answer the below survey questions honestly. Your participation or non-participation will not affect your grade or class standing. These surveys are anonymous. Participation is voluntary, and you can choose to not answer any question that you do not want to answer, and you can stop at anytime.

INSTRUCTIONS: Please respond to the statements below using the following scale 1-5 scale. **5-strongly agree, 4-Somewhat Agree, 3-Neutral, 2-Somewhat Disagree, 1-Strongly Disagree.**

1. I am confident in my ability to connect specific evidence to reasoning and claims made in science class.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

2. I am successful at constructing and utilizing evidence-based claims in my writing.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

3. Self-Assessments are helpful in improving my understanding of content covered in science.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

4. Self-Assessments are helpful in improving my writing skills in science.

    1   2   3   4   5   6

    Strongly Agree   Strongly Disagree

5. The one-minute essay helped me better understand material I've learned in science.

    1   2   3   4   5

    Strongly Agree   Strongly Disagree

6. I have a good idea of what I know and what I need to work on in science class.

   1   2   3   4   5

   Strongly Agree   Strongly Disagree

7. The use the one-minute essay helped me better connect evidence with claims made in class.

   1   2   3   4   5

   Strongly Agree   Strongly Disagree

8. I am comfortable and confident evaluating my own work by means of a self-assessment

   1   2   3   4   5

   Strongly Agree   Strongly Disagree

9. Using the self-evaluation form helped me *identify* some of weaknesses in writing.

   1   2   3   4   5

   Strongly Agree   Strongly Disagree

10. Self-evaluation use has improved my comprehension of content covered in this class

   1   2   3   4   5

   Strongly Agree   Strongly Disagree

11. After utilizing the self-assessment rubric, I feel like I was able to improve my writing outputs in this class.

   1   2   3   4   5

   Strongly Agree   Strongly Disagree

APPENDIX F

SEMI-STRUCTURED STUDENT INTERVIEW QUESTIONS

Semi-Structured Student Survey Questions

Students were interviewed informally in groups of two after the treatment period was completed. Prior to asking the below questions I reminded the students that their answers in this setting had no merit on their grades. I told them they did not have to answer any question they did not feel comfortable with. I also told them their answers, regardless of whether they felt positively or negatively about the experience, would only help me going forward in teaching 8th grade science.

1. How important do you feel it is to use specific evidence to back claims made in science class?

2. How effectively do you feel you can do this in writing?

3. Do you like using self-assessments?

4. Do you find them helpful?

5. Does using pneumatic device, like C.E.R help remind you of the importance of connecting evidence to reasoning and claims in your science writing?