2010

Is a student's conceptual knowledge affected by their vocabulary instruction?

Daniel C. Cox
Louisiana State University and Agricultural and Mechanical College, dcox1@ebrschools.org

IS A STUDENT'S CONCEPTUAL KNOWLEDGE AFFECTED BY THEIR VOCABULARY INSTRUCTION?

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Natural Sciences

in

The Interdepartmental Program in Natural Sciences

by
Daniel C. Cox
B.S., Louisiana State University, 2006
August 2010

Acknowledgements

I would like to thank all of the faculty and fellow students in the masters of natural science program at LSU who helped me, knowingly or otherwise, to complete my thesis. I would also like to thank my wife and friends for their support and encouragement throughout the process of researching and creating this thesis.

Table of Contents

List of Tables

List of Figures

Abstract

The following paper explores the possible relationship between the manner in which a student is taught vocabulary in a high school physics classroom and the ability of those students to correctly answer physics conceptual questions. Two high school physics classes taught by the same teacher (eight or nine students in one and seventeen in the other) received identical instruction throughout multiple units with the exception of vocabulary instruction. Differences were observed in the students' ability to correctly answer conceptual questions in only one topic, heat. The students did not get better at answering a question concerning a phase change or another question that the direct instruction group already knew (80% of the students got the right answer on the pretest and kept that same answer on the posttest).

Introduction

As long ago as 1917, questions were being raised (Thorndike) about the intricacies of understanding phrases, sentences, and paragraphs. Students were allowed to read a paragraph and then asked simple questions (short answer format) whose answers could be determined using the assigned reading. Not only did a large number of students answer incorrectly, but the types of incorrect answers given were difficult for the researcher to classify.

In my school district, teachers are repeatedly asked to integrate general literacy activities into their instruction. Types of activities vary, but they share the common goal of improving a student's ability to read and write. Despite the intuitively obvious benefits of a policy meant to enhance literacy skills, it would be much easier to enforce such a policy if teachers felt that there was a carry-over effect that increased student performance in other areas. Fortunately, one particular survey of literature exists (Stahl & Fairbanks 1986) that suggests a carry-over effect into the ability of a student to understand passages involving untaught vocabulary. It seems that students can receive vocabulary instruction for a particular set of words and improve not only their ability to comprehend passages involving those words, but also their ability to comprehend passages that don't include the taught vocabulary. This conclusion was made after the paper's authors compiled the results of roughly 120 different sources of information. While it is nice that an increase in general reading ability can occur, is there a similar increase in the ability to correctly answer conceptual questions in a physics setting?

As a physics teacher, I wanted to know if time spent on literacy instruction would have any positive influence on my students' scientific literacy. That led me to implement the most basic form of literacy instruction, vocabulary instruction, and to measure a fundamental ability in a physics student, the ability to correctly answer a conceptual question.

Literature Review

Textbooks

The textbooks in classrooms can be very difficult for students to read, much less comprehend. The vocabulary within textbooks has been analyzed on more than one occasion and it has been discovered that there are multiple roadblocks to reading comprehension. Students must deal with technical vocabulary that would only appear in a science textbook, which can lead to a coping mechanism of attempting to memorize as much of the information as possible without trying to understanding it (Groves 1995, Songer & Linn 1991). Others have found that a student must also understand the "procedural" vocabulary and non-technical words in science textbooks in order to comprehend what they are reading (Marshall & Gilmour 1991, Marco 1999). This means that there are common words and phrases that teachers assume are known to his or her students which are, in reality, troublesome for them. Recommendations have been made by researchers to teach using vocabulary as a stepping stone into conceptual discussions in order to avoid rote memorization of facts. Teachers are also asked to integrate the teaching of nontechnical and procedural vocabulary into their lessons since a superficial understanding of these words can negatively affect reading comprehension (Groves 1995, Songer & Linn 1991, Marshall & Gilmour 1991, Marco 1999).

There are, however, researchers who feel that science textbooks get a bad reputation because words that are typically found in science textbooks are not included on word lists used to rate their readability (O'Toole & Bedford 1969). In an attempt to rectify the situation, words commonly found in six different textbook series used in elementary grades were compiled. These words were then used to construct a multiple-choice examination in order to determine which words were known by most of the students in a group of 720 fourth graders. All words that 60% or more of the students knew were compiled into a more refined list and this list was used in addition to a standard readability list in order to reevaluate science books not used in the initial gathering of words. All five of the books assessed dropped in reading level by at least one grade. A study like this seems to refute the argument that the terminology in science books is difficult to understand, but what is more likely is that it illustrates the need for readability lists to be updated.

Vocabulary Instruction

Research has explored the effect of different vocabulary instructional techniques on vocabulary retention and reading comprehension skills. The common practice of assigning vocabulary to define was denounced in a paper that found the activity to be “meaningless” (Thelen 1986). Copying definition was not found to increase reading comprehension. The study suggested providing students with not only the meaning of words, but the reason why they should care about their definition. They also feel that building on students’ current understanding (schema) will lead to better reading comprehension.

Research has also established a connection between student participation and long-term vocabulary retention in groups of fifth grade students (Stahl & Clark 1987). Students in this study were split into three groups and given instructions concerning a discussion in class. One group was told that they would be called on to participate in the discussion, another that they might be called on, and the final group was told that they were simply allowed to listen. Class discussion that involved any level of student participation led students to retain vocabulary better than the students who were told to learn by listening. Interestingly, it did not seem to matter whether the students did or did not actually participate, only that they thought they could be called on to participate. In the short term, there was no significant difference between the performance of the three groups, but an assessment given a couple weeks after the first one showed a difference in retention between the discussion groups and the listening groups.

Other researchers have made similar conclusions by employing different techniques that also seek to engage students in learning (instead of turning them into spectators). One study used college freshmen in a physical science class as its subjects (making its subjects similar to the ones used in this study) (Snouffer & Thistlewaite 1979). They employed two strategies, graphic pre-organizers and pre-reading vocabulary instruction, in an attempt to increase the students’ ability to understand and remember what they read. The graphic organizers were similar to a chapter outline or a concept map and the pre-reading vocabulary instruction introduced the students to eight words that students would encounter in a reading selection. They found that science reading comprehension and the ability to recall information was significantly improved by pre-reading tasks, with pre-organizers the clear winner. The researchers admit that the reader should approach their results with caution since pre-organizers had not shown such promise in other studies.

Another study found success in the use of an activity dubbed *“Possible Sentences* (Stahl & Kapinus 1991).” Students would use definitions of words, both familiar and unfamiliar, to construct sentences that would be analyzed for correctness after reading an assignment. The sentences that were incorrect (vocabulary improperly used) were corrected by the students after the activity in the form of a class discussion. Students were forced to participate more fully in these activities than simple reading activities. It is believed that this deeper participation aided the students in learning vocabulary and recalling factual information.

Further reinforcement comes from a study performed on fourth and eighth grade science students who were taught using discussion and hands-on activities in a class where inquiry activities dominated instruction (Carlisle, Fleming, & Gudbrandsen 2000). Significant increases were found in these students' knowledge of topical science vocabulary. Off-topic science vocabulary was included in the assessments and no gain occurred. They did, however, measure an increased performance on a test of "applied problems," which may have implications for this thesis' research. This study also determined that the more vocabulary knowledge a student possesses, the more they will learn and vice versa.

Another paper establishes the usefulness of hands-on activities by comparing them to instruction involving no direct vocabulary instruction and glossary look-up accompanied with using words in a sentence (Lloyd & Contreras 1987). In addition to the fourth grade students in the study scoring higher on an assessment of vocabulary knowledge, there was another interesting result. A control group that received no instruction and the group that used the glossary and sentence-writing method had indistinguishable gains in knowledge.

It should be mentioned, however, that not all researchers have had success with pre-teaching strategies. One study involving 6th grade science students found no difference between the effectiveness of three different pre-reading strategies: traditional glossary usage, determining definitions from context, and determining the characteristics of words with a semantic map (Seaver 1991). Each group's reading comprehension and vocabulary mastery was indistinguishable from the others.

In a research review that summarized techniques used to teach vocabulary effectively in multiple disciplines, recommendation were made to science teachers for better vocabulary instruction (Harmon, Hedrick, & Wood 2005). Providing more reading opportunities, teaching vocabulary explicitly, using books that are below grade level, and other tactics are cited as effective means of teaching content-specific vocabulary.

Some researchers have found situations that seem to enhance students' abilities to perform in both science and language when the science curriculum is allowed to influence reading instruction (Romance Vitale 1992). Students whose English classes were essentially combined with their science classes performed better on Iowa Test of Basic Skills reading subtest and on the Metropolitan Achievement Test science subtest. Incidentally, they displayed a more positive attitude towards science and reading.

Also worth mentioning are the effects of changes in vocabulary instructional techniques on ESL students. While my students' primary language is English, I suspect that they might feel otherwise when attempting to read science text. One study involved sixth grade migrant students (Moran 1990). The teacher in this study was having difficulty teaching the science curriculum prescribed by her district and she felt it was due to the fact that not all of her students possessed an appropriately extensive vocabulary. Altering the vocabulary instruction of these students had a pronounced effect on their class performance. The students' class grades increased and their gain scores on a pretest / posttest.

Conceptual Understanding

Some research states that it is probably not necessary to understand terminology or possess above average reading comprehension skills to succeed in understanding science concepts. If this is true, then this study is not likely to discover a link between vocabulary instruction and conceptual development. A paper from more than four decades ago describes a researcher's attempt to create an assessment for children that uses absolutely no words (Haney 1965). It is a test meant to assess a child's understanding of animals. The fact that the researcher was attempting to do so implies that they believe in a lack of correlation between knowledge of words and knowledge of scientific concepts.

More recent research has been performed involving a group of fifth graders being taught the process of photosynthesis (Brown & Ryoo 2008). Some of the students were taught vocabulary relevant to the lesson followed by lessons intended to explain concepts. Other students were taught in the reverse order and everyday words were used in place of technical ones until they were eventually taught. It would be interesting enough if the research discovered no difference between the two groups, but they actually found that the group who was taught using their "content-first" approach performed better than the other group.

Another article describes the very low correlation that seems to exist between reading comprehension and problem-solving skills in a group of physics students in Israel (Koch & Eckstein 1995). In this study, researchers deal with junior college students. Students' reading comprehension abilities were evaluated using assessments that required very little problem-solving ability. Also, their problem-solving abilities were assessed using questions that were easy to comprehend. The study determined that the two abilities are virtually independent of each other. One intriguing bit of their study concerned students who could read a passage and then, after completing the reading, correctly identify information that was not in the passage. The students who were better at performing this task performed somewhat better at the problem-solving activities. In their conclusion, they decided that in problem-solving it is important to know not only what is given to you, but also what is not given to you.

Not all research in this area has served to establish a lack of correlation, however. A 1972 study involving undergraduate students in an educational psychology course shows that similar groups of students have different levels of understanding of the conceptual meaning of a word when they receive different forms of instruction (Anderson & Kulhavy 1972). One group was asked to read a word's definition aloud three times and the other group was asked to create sentences using those same words. Students then took a test where they were asked to choose the sentence that best demonstrated the word's meaning. The researchers were careful not to use the word itself or any substantial words from the word's definition when creating sentences for the test. Also, the test questions were all norm-referenced against students in the same class that received no instruction of any sort. Ultimately, those who created sentences significantly outperformed those who repeated definitions aloud.

In order to assess students' conceptual knowledge, some researchers have developed assessment devices that are now widely accepted as valid. One of these is the Force Concept Inventory (FCI) (Hestenes, Wells, & Swackhamer 1992). It has been effectively used on high school students and is meant to detect common misconceptions held by student in introductory physics courses. The questions on the exam are multiple-choice, but are not easily answered unless the test-taker understands the concept involved. The incorrect answers are carefully tailored to identify why students didn't know the correct answer.

Another assessment tool similar to the FCI behaves as more of a measure of instruction. The Mechanics Baseline Test (MBT) does not seek to identify physics misconceptions (Hestenes & Wells 1992). Instead, it tries to assess whether or not the concept was learned.

A third tool, the Force and Motion Concept Evaluation (FMCE) was administered to introductory physics students in college and helped establish the lack of effectiveness of "traditional" teaching methods in physics (Thornton & Sokoloff 1998). In contrast, it helped show that a hands-on, more interactive approach yields better understanding of physics concepts.

This study also made use of practice problems for AP physics tests. Some of these questions are accessible to the public and some are only accessible to teachers of AP courses. These exams are designed to evaluate whether or not a high school student should be given college credit for introductory college physics and have been in use for decades (AP Central).

Synopsis

In summation, the literature available suggests that the more a student is immersed in an activity, the more he or she will learn from it. It would seem that Piaget (1952) was right in his belief that a student learns best through personal experiences. Success has primarily been found with activities involving hands-on participation, discussion, and the general application of something more than superficial effort. When these sorts of changes are imposed upon vocabulary instruction, one can readily expect that students' ability to comprehend what they read will improve. A connection between reading comprehension and higher-order think abilities is lacking, though. In fact, there is convincing evidence that the two skills are disconnected and that concepts can be taught without the use of technical vocabulary. While it seems that some increase in ability will come from direct vocabulary instruction when compared to indirect instruction, it does not seem likely that the benefit will manifest itself in a student's ability to correctly answer conceptual questions in a physics setting.

Materials and Methods

Subjects

The groups in this study were classes of high school physics students at a public school in Louisiana. The students in the groups ranged in age from sixteen to twenty years in age, but all were classified as seniors by the school. One of the groups in this study was composed of eight to nine students (one dropped at mid-year) and the other group was composed of seventeen students. Both classes met in the morning. The smaller class took place between seven and eight o'clock and the larger of the two took place between ten and eleven o'clock. Both classes were of mixed gender and race. Both of the physics classes were classified by the school as "regular" classes. This means that the subjects all opted out of taking the Advanced Placement (AP) version of the physics course. Table and figure one show the differences in LEAP scores for the two groups. The standard error in the table refers to the error in the average scaled score. The degree of freedom (df) has to do with the number of subjects involved in the comparison. Lastly, p-values indicate whether or not it's likely that my comparison is made with an incorrect assumption (null hypothesis). All comparisons in this study assume that the differences between two averages are zero and a p-value less than 0.05 will indicate a flawed null hypothesis. Standardized test scores for the students in one group were similar to those of the students in the other.

Table 1 – LEAP Score Comparisons

	Scaled Score Difference	Standard Error	t (null: difference=0)	df	p-value
English	8	11	0.72	12	0.48
Science	15	16	0.97	9	0.36
Math	13	20	0.66	8	0.53

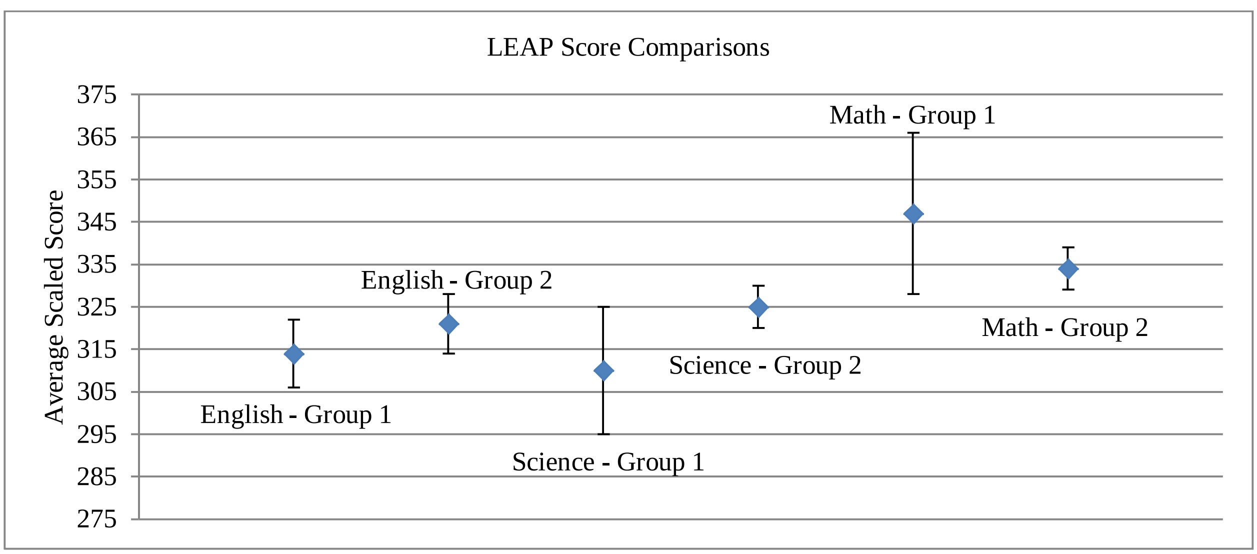

Figure 1 – LEAP Score Comparisons

Instruction

The two groups received differing vocabulary instruction. The two methods employed were direct and indirect vocabulary instruction. Both techniques used scanned passages from a Holt Physics textbook, but the implementation differed slightly. The direct instruction class was given ten minutes to paraphrase definitions for bold words located in the passages. The glossary definitions were located in the margins of the reading assignment. A class discussion of individual student responses followed the paraphrasing activity. The discussion lasted five minutes. During this discussion, students were asked to comment on each other's work and the teacher gave general advice to students wishing to improve their responses. The indirect instruction group was asked, instead, to silently read the textbook passages for fifteen minutes. In order to provide similar amounts of motivation, students were assigned participation grades for taking part in the vocabulary instruction. The grade received for full participation in each group carried the same weight. The treatment was applied one time for each unit of instruction involved in this study. In an attempt to control any effects due to class size, the group that received direct vocabulary instruction for one topic invariably became the control group (indirect instruction) for the next topic and vice versa. This process continued throughout the experiment.

Assessment

Pretests were administered at the beginning of each unit. The questions for these tests were selected from established assessment devices and were only modified in the sense that they had been separated from the rest of the questions in their original assessment device. Each of the six pretests consisted of three of these questions except for the momentum pretest (two questions) and the heat pretest (four questions). Within the first couple of days of instruction the classes received their vocabulary instruction. At the end of each unit's instruction, a posttest that was identical to the pretest was administered. Students were given as much time as needed to answer questions on the pretests and posttests. At no point were the answers to any pretest or posttest discussed in class and all other assignments took place for similar amounts of time on similar days. If either class was disrupted by a large field trip or assembly, then the other class was allowed free time to study for other classes in order to compensate for the difference in instruction time. Also, no aspect of regular instruction was tailored in order to "teach to the test." It was as though the teacher forgot about the existence of the pretests and posttests until it was time to administer them. Each class received identical participation grades for taking the pretests and posttests in an attempt to equalize motivation.

Topics

Students were acclimated to the pretest-posttest procedure throughout the beginning of the school year (pre-November). Three pretests and posttests were administered during that time. Results from pretests and posttests for kinematics and Newton's laws of motion were not recorded for this reason. The first unit of instruction for which data was recorded dealt with circular motion (4 instructional days). After that, angular kinematics (5 instructional days), conservation of energy (13 instructional days), conservation of momentum (8 instructional days), heat transfer & calorimetry (16 instructional days), and simple harmonic motion (6 instructional days) was taught. Each unit's instruction was generally similar. The beginning of each unit involved a day or two of note-taking and discussion. Following that, students were shown how to solve word problems and given opportunities to do so on their own when they were comfortable enough to do so. Further class discussion and lecture occurred when needed (as determined by teacher). The tests for each unit were made from problem-solving questions that were similar to those practiced in class. The pretest and vocabulary instruction required 20-25 minutes and the post test required 5-10 minutes.

Results and Discussion

Student responses to pretests and posttests were recorded in a spreadsheet and numbers of correct responses were tallied for the pretests and posttests. Calculated gains are normalized. This means that, instead of merely subtracting pretest scores from posttest scores (a method that produces the raw gain), the realized gain has been divided by the possible gain to get a ratio known as the normalized gain. Obviously this value will, at most, be equal to 1. This causes a problem when a subject scores perfectly on a pretest. A perfect pretest score leads to a possible gain of zero and a normalized gain with a denominator of zero. Only one student received a perfect score on a pretest so a negligible amount of data was lost due to this effect.

The Welch t-test was used in all t-test analysis. It is similar to a paired t-test, but assumes that the two sample means have a different variance (s^2). What is noteworthy about the test is that the standard error in the mean (which ultimately affects the t-value) and effective degrees of freedom are dependent on the variance of both samples. Also, the test's assumption of normality need not be true in order to establish whether or not two means are similar. The formulas used to calculate the standard error and the degrees of freedom are located in appendix A of this thesis. The t-value and degrees of freedom obtained when using this test can be used to obtain a p-value (the probability that rejecting the null hypothesis would be incorrect). All significance tests in this thesis were performed with an alpha of 0.05. This means that a given null hypothesis was rejected only if the p-value was less than 0.05.

In order to determine whether or not a bias existed in the ability of the two groups' to answer the questions on the pretests, a comparison of pretest scores was performed for each topic. In an attempt to increase group homogeneity, students who were not present each day of instruction were not included in this analysis. This method will continue to be used unless stated otherwise. If one group had an advantage over the other, the difference in pretest scores would be something other than zero (null hypothesis: group 1 – group 2 = zero). The analysis is displayed in table 2 and figure 2:

Table 2 - Group Bias Analysis

	Difference (Group 1 – Group 2)	Standard Error of Difference	t (null: Group 1-Group 2=0)	df	p (two tail)
Circular	0.31	0.34	0.91	14	0.38
Angular	-0.43	0.26	1.7	13	0.12
Energy	0.17	0.37	0.45	8	0.67
Momentum	0.18	0.29	0.62	6	0.56
Heat	0.2	0.37	0.53	6	0.61
SHM	0.37	0.33	1.1	8	0.29

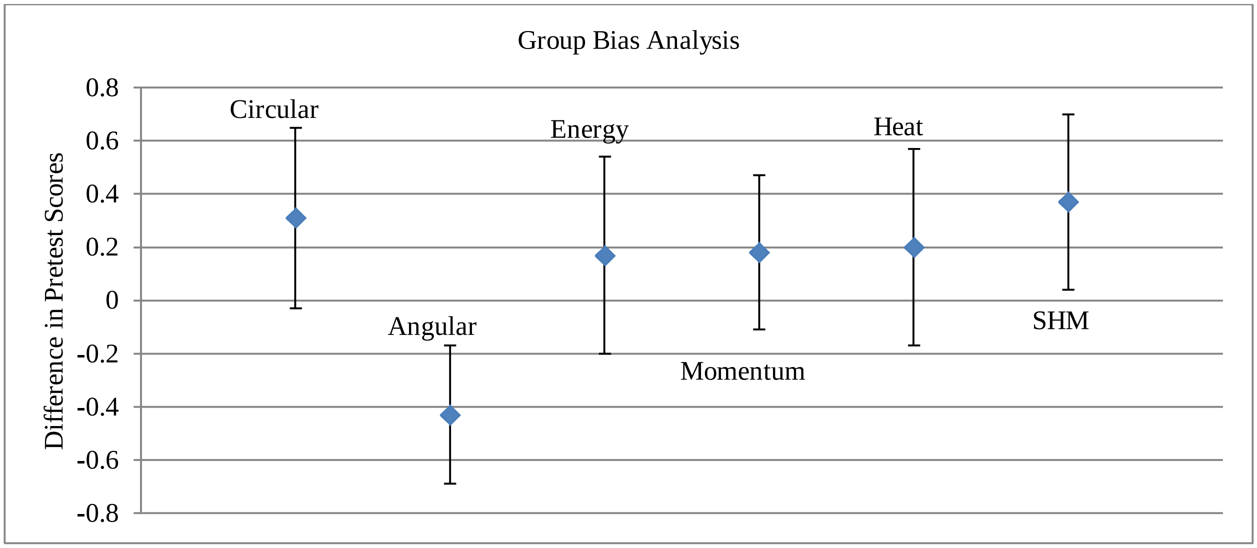

Figure 2 – Group Bias Analysis

No bias can be claimed to exist for any test since each p-value was larger than 0.05. Next, the average normalized gains for each group were compared and results displayed in table 3 and figure 3:

Table 3 – Change in Normalized Gain Scores for Each Unit

	Mean Difference	Standard Error	t (null: difference=0)	df	p
Circular Motion	-0.10	0.18	0.55	8	0.30
Angular Kinematics	-0.09	0.23	0.39	12	0.35
Energy	-0.06	0.37	0.15	9	0.44
Momentum	-0.07	0.30	0.23	7	0.41
Heat	0.48	0.14	3.6	7	0.0046
SHM	-0.07	0.22	0.30	7	0.39

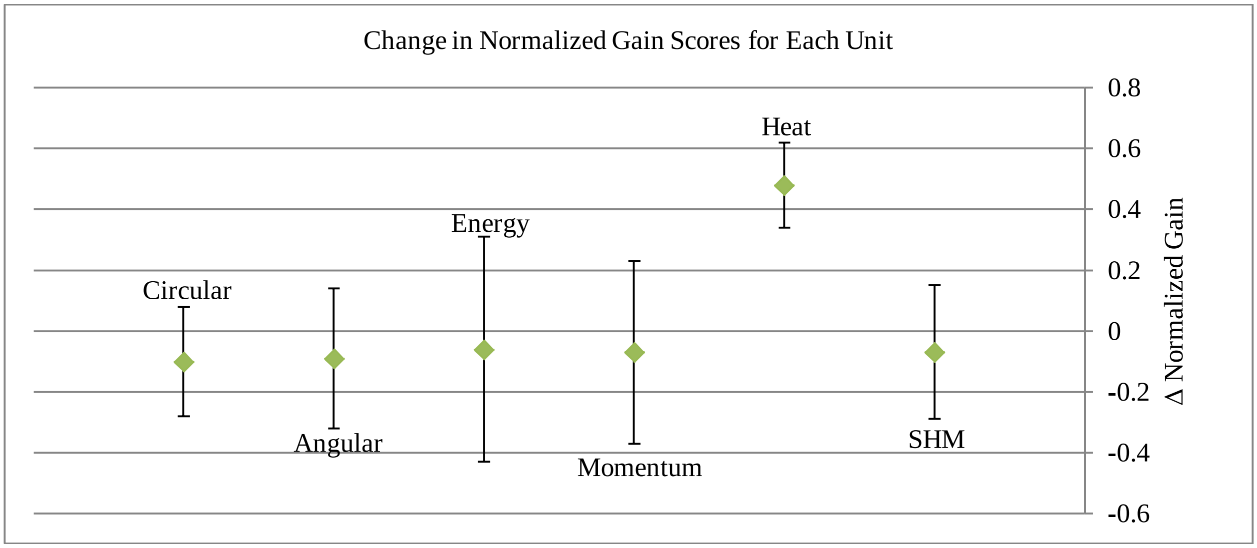

Figure 3 – Change in Normalized Gain Scores for Each Unit

The mean difference for simple harmonic motion, momentum, energy, angular motion, and circular motion are all consistent with zero since the p-values for those topics are all larger than 0.05. The heat comparison does, on the other hand, show a significant gain with a p-value less than 0.05. In an attempt to better understand the significant difference in gains during the heat unit, gains for individual pretests / posttests were analyzed. The results for direct instruction pretests / posttests are in table and figure 4 and indirect instruction results are in table and figure 5:

Table 4 - Normalized Gains for Direct Instruction

	Mean Gain	Standard Error	t (null: gain=0)	df	p
Circular Motion	-0.08	0.17	0.50	7	0.32
Angular Kinematics	-0.14	0.13	1.7	12	0.056
Energy	0.00	0.29	0	5	0.50
Momentum	-0.17	0.17	1.0	8	0.17
Heat	0.483	0.085	5.7	4	0.0024
SHM	-0.067	0.079	0.84	9	0.21

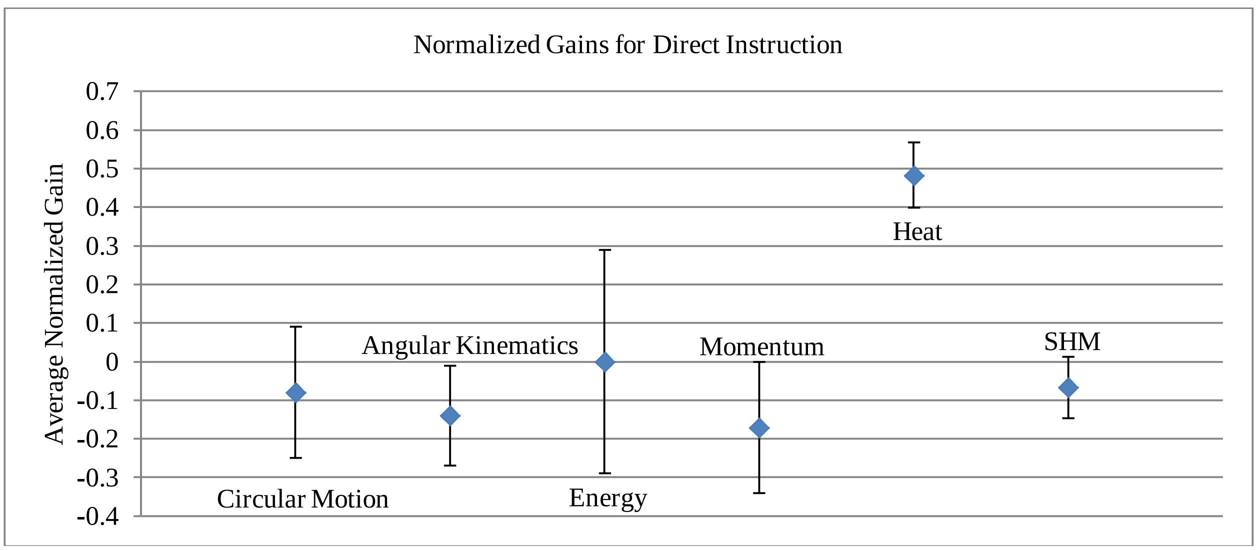

Figure 4 - Normalized Gains for Direct Instruction

Table 5 - Normalized Gains for Indirect Instruction

	Mean Gain	Standard Error	t (null: gain=0)	df	p
Circular Motion	0.013	0.055	0.23	12	0.41
Angular Kinematics	-0.23	0.18	0.78	6	0.23
Energy	0.06	0.23	0.24	5	0.41
Momentum	-0.10	0.24	0.41	4	0.35
Heat	0.00	0.11	0	4	0.50
SHM	0.00	0.21	0	6	0.50

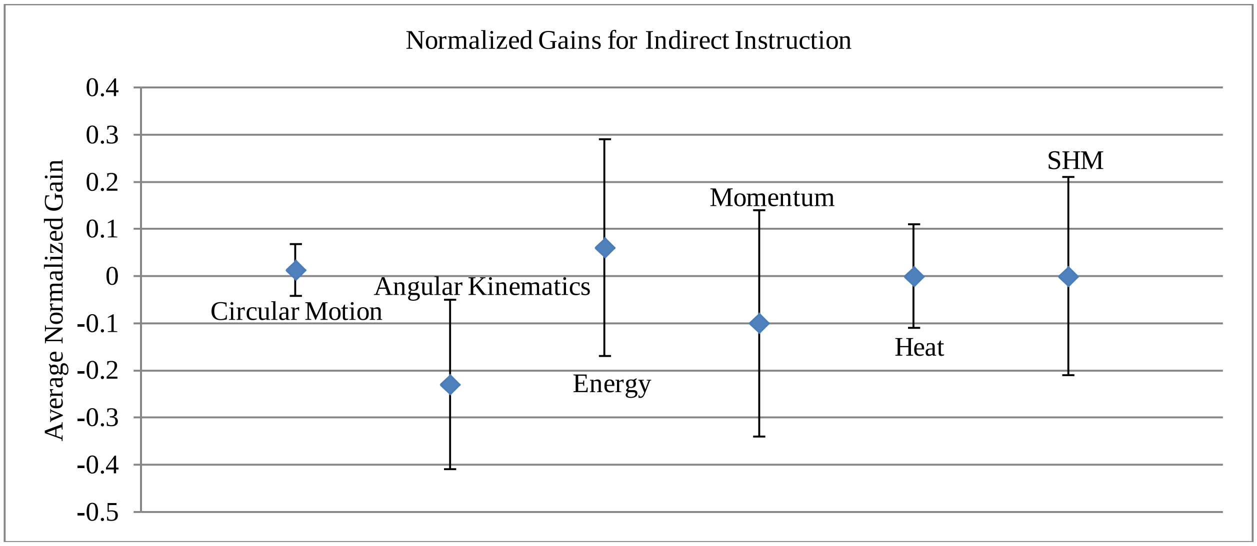

Figure 5 - Normalized Gains for Indirect Instruction

All p-values for the normalized gains in indirect instruction are larger than 0.05 and cannot be distinguished from zero. Within the direct instruction groups, all of the p-values were larger than 0.05 except for the heat unit. Its p-value was below 0.05. This result supports the earlier analysis showing a difference in gains between groups.

Differences in raw gains for individual questions are compared below in an attempt to find out what factors led to the gain on the heat test for direct instruction students. Raw gains were used since any correct scores on the pretest questions would lead to invalid normalized gains (roughly half of the data would be lost). Results are in the following table:

Table 6 - Individual Question Raw Gain Analysis

	Gain Differences	Standard Error	t (null: direct – indirect > 0)	df	p (one-tail)
Circular – Question 1	-0.33	0.18	1.8	9	0.052
Circular – Question 2	0	0.25	0	15	0.50
Circular – Question 3	0.25	0.16	1.5	7	0.085
Angular – Question 1	-0.23	0.35	0.70	7	0.25
Angular – Question 2	-0.08	0.10	1.0	12	0.17
Angular – Question 3	0.14	0.34	0.47	10	0.33
Energy – Question 1	0	0.37	0	7	0.50
Energy – Question 2	-0.17	0.20	1.0	5	0.18
Energy – Question 3	0.17	0.43	0.45	8	0.33
Momentum – Question 1	0	0	-	-	-
Momentum – Question 2	-0.11	0.41	0.30	7	0.39
Heat – Question 1	0.20	0.20	1.0	4	0.19
Heat – Question 2	0	0.28	0	8	0.50
Heat – Question 3	0.60	0.24	2.4	4	0.035
Heat – Question 4	0.60	0.24	2.4	4	0.035
SHM – Question 1	0.10	0.24	0.42	8	0.34
SHM – Question 2	-0.19	0.30	0.61	7	0.28
SHM – Question 3	-0.10	0.10	1	9	0.17

Each individual question's p-value was greater than 0.05 except for questions three and four on the heat test. Student responses to question one on the momentum test didn't change at all and no student correctly answered the question. Since this leads to a standard error of zero, the t-value was impossible to calculate (t-values are inversely proportional to standard error).

A quick look at the questions on the heat test will reveal reasons for the lack of raw score gains on questions one and two. Pretest scores for question one show that four of the five students whose responses were used in the direct instruction data were correct to begin with and didn't change on the posttest. It is difficult to measure a gain if students already know the answer. Question two was the only question on the test dealing with a phase change while the rest of the questions dealt only with temperature changes.

Conclusion

According to the data analysis in the previous section, the type of vocabulary instruction a student receives does not generally enhance their ability to understand physics concepts. The study does suggest, however, that direct vocabulary instruction had a positive impact on the subjects' ability to correctly answer conceptual questions on the heat test. The two groups taking the heat test performed similarly on their heat pretests, but the direct instruction group had a non-zero normalized gain. Subjects receiving indirect instruction, on the other hand, achieved a normalized gain on the same test that was indistinguishable from zero. Analysis of individual question raw gain score differences shows positive gains on two of the four questions on the heat test. The two questions on the heat test that didn't show a difference in gain scores were either too easy (the students already knew the answer) or they dealt with a concept that was different from the concepts in the two questions that showed a gain. All of these things support the possibility that students benefitted from direct vocabulary instruction within this topic.

There are, however, factors that cast doubt on the positive gains and differences in gain scores. The sample sizes used in this study make results difficult to trust for the same reason that a possible outcome is much more likely to occur five times in a row than it is to occur 100 times in a row. The possibility that the positive gains are due to chance can only be proven or ruled out with more experimentation or experimentation with more subjects. Also, the amount of instructional time for the heat unit was longer than any other. The longer instructional time allows students to, by chance, have life experiences that could affect their posttest scores. This study's small sample size exacerbates this potential effect. A couple of students in a group of 100 who perform better due to some enlightening life experiences will have a much smaller effect on average gain scores than in a group of five. Also, the large amount of time between pretest and posttest allows the teacher more opportunities to unwittingly make mistakes when attempting to administer non-vocabulary instruction in identical ways. Lastly, the results for question #1 on the momentum pretest / posttest suggest that the question is poorly worded or misleading. With eighteen total questions, it seems likely that other questions could be flawed in some way that data analysis didn't make obvious. Any future experiments should decrease the amount of time between pretests and posttests and work with larger sample sizes to rule out such possibilities. Extreme care should also be taken to ensure that the assessment devices used are as reliable as possible.

References

Anderson, R. C., Kulhavy, R. W. (1972) Learning Concepts from Definitions. *American Educational Research Journal*, 9, 385-390

AP Central. 2010. The College Board. <http://apcentral.collegeboard.com/apc/Controller.jpf>

Brown, B. A., Ryoo, K. (2008). Teaching Science as a Language: A ''Content-First'' Approach to Science Teaching. *Journal of Research in Science Teaching*, 45(5), 529-553.

Carlisle, J. F., Fleming, J. E., & Gudbrandsen, B. (2000). Incidental word learning in science classes. *Contemporary Educational Psychology*, 25(2), 184–211.

Dimitrov, D. M., Rumrill, P. D. (2003) Pretest-posttest designs and measurement of change. *Work: A Journal of Prevention, Assessment, and Rehabilitation*, 20(2), 159-165.

Groves, F. H. (1995). Science vocabulary load of selected secondary science textbooks. *School Science and Mathematics*, 95(5), 231–235.

Haney, R. E. (1965). The Development of a Non-Verbal Test of Children's' Concepts of Animals. *Journal of Research in Science Teaching*, 3, 198-203.

Harmon, J. M., Hedrick, W. B., Wood, K. D. (2005) Research on Vocabulary Instruction in the Content Areas: Implications for Struggling Readers. *Reading and Writing Quarterly*, 21, 261-280.

Hestenes, D., Wells, M. (1992) A Mechanics Baseline Test. *The Physics Teacher*, 30, 159-166.

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

Koch, A., Eckstein, S. G. (1995) Skills Needed for Reading Comprehension of Physics Texts and Their Relation to Problem-Solving Ability. *Journal of Research in Science Teaching*, 32(6), 613-628.

Lloyd, C., Contreras, N. (1987) What Research Says: Science Inside-Out. *Science and Children*, 25(2), 30-31.

Marco, M. (1999). Procedural vocabulary: Lexical signaling of conceptual relations in discourse. *Applied Linguistics*, 20(1), 1–21.

Marshall, S. & Gilmour, M. (1991). Words that matter in science and technology: A study of Papua New Guinean students' comprehension of nontechnical words used in science and technology. *Research in Science and Technological Education*, 9(1), 5–16.

Moran, M. Q. (1990). Improving science vocabulary of sixth grade migrant students through a program of intensive direct vocabulary instruction. Eric No. ED 325 301.

O'Toole, R. J., Bedford, J. P. (1969) Science Vocabulary and Readability Level. *Journal of Research in Science Teaching*, 6, 161-162.

Piaget, J. (1952) "The Origins of Intelligence in Children." New York: International Universities Press.

Seaver, J. (1991). The effect of pre-reading vocabulary development strategies on sixth grade science text results. Eric No. ED 329 952.

Snouffer, N. K. & Thistlethwaite, L. L. (1979). The effects of the structured overview and vocabulary pre-teaching upon comprehension levels of college freshmen reading physical science and history materials. Paper presented at the 29th Annual Meeting of the National Reading Conference, San Antonio, TX, Nov. 29–Dec. 1. Eric No. ED 182 728.

Songer, N. B. & Linn, M. C. (1991). How do students' views of science influence knowledge integration? *Journal of Research in Science Teaching*, 28, 761–784.

Stahl, S. A. & Clark, C. H. (1987). The effects of participatory expectations in classroom discussion on the learning of science vocabulary. *American Educational Research Journal*, 24(4), 541–555

Stahl, S. & Fairbanks, M. (1986). The effects of vocabulary instruction: A model based meta-analysis. *Review of Educational Research*, 56(1), 72–110.

Stahl, S. A. & Kapinus, B. A. (1991). Possible sentences: Predicting word meanings to teach content area vocabulary. *The Reading Teacher*, 45(1), 36–43.

Thelen, J.N. (1986). Vocabulary instruction and meaningful learning. *Journal of Reading*, 29(7), 603-609.

Thorndike, E. L. (1917). Reading as reasoning. *Journal of Educational Psychology*, 8, 512- 518.

Thornton, R. K., Sokoloff, D. R. (1997). Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula. *American Journal of Physics*, 66(4), 338-352.

Appendix – Welch t-Test Calculations

Calculation of standard error for two samples:

$$SE_c = \sqrt[2]{(SE_1)^2 + (SE_2)^2}$$

Calculation of effective degrees of freedom:

$$df = \frac{((SE_1)^2 + (SE_2)^2)^2}{\frac{(SE_1)^4}{n_1 - 1} + \frac{(SE_2)^4}{n_2 - 1}}$$

SE_c = the combined standard error of the two samples

SE_1 = the standard error of sample one

SE_2 = the standard error of sample two

df = effective degrees of freedom

n_1 = number of subjects in sample one

n_2 = number of subjects in sample two

Vita

Daniel Cox was born in Louisiana, in 1983. He attended public schools in the city of Hammond, Louisiana, and graduated as the salutatorian from Hammond High School in 2001. Within the same year, he enrolled in Louisiana State University in Baton Rouge. Originally a mechanical engineering student, he switched his major during his junior year to physics with a concentration in secondary education. In 2006, he received a Bachelor of Science in physics and began teaching physics at the high school level in Baton Rouge at Broadmoor High School. He is also certified to teach high school mathematics in addition to the physical sciences in the state of Louisiana.