THE EFFECTS OF VOCABULARY INTERVENTION ON NINTH GRADERS’ UNDERSTANDING OF PLATE TECTONICS

A thesis presented to
the faculty of
the College of Arts and Sciences of Ohio University

In partial fulfillment
of the requirements for the degree
Master of Science

Timothy J. Sekula
November 2006
This thesis entitled
THE EFFECTS OF VOCABULARY INTERVENTION ON NINTH GRADERS’ UNDERSTANDING OF PLATE TECTONICS

by
TIMOTHY J. SEKULA

has been approved for
the Department of Geological Sciences
and the College of Arts and Sciences by

Dina L. Lopez
Associate Professor of Geological Sciences

Benjamin M. Ogles
Dean, College of Arts and Sciences
Abstract

SEKULA, TIMOTHY J., M.S., November 2006, Geological Sciences

THE EFFECTS OF VOCABULARY INTERVENTION ON NINTH GRADERS’ UNDERSTANDING OF PLATE TECTONICS (87 pp.)

Director of Thesis: Dina L. Lopez

Integration of vocabulary tools into content specific curriculums may ultimately provide students and teachers with the opportunities to address the overwhelming breadth of vocabulary that exists in the Ohio Academic Content Standards (OACS). In an effort to overcome the language barrier that exists in science curriculums, students and teachers need to be able to spend time discovering meaningful applications and strong connections to content through the establishment of “deep word knowledge” rather than struggling with misconceptions and inadequate comprehension due to mandated state curriculums. This study investigated the possible benefits provided by language interventions in a secondary science setting with the goal of determining acceptable and attainable teaching methods capable of covering state academic content standards laden with large amounts of new vocabulary. The vocabulary interventions utilized with the experimental group demonstrated effectiveness for increasing student comprehension rates as seen through quantitative analysis of pre- and post test assessment scores.

Approved:

Dina L. Lopez
Associate Professor of Geological Sciences
Acknowledgments

Co-Advisor
Dina Lopez, Ph. D.

Co-Advisor
Julie Libarkin, Ph. D.

Committee:
Julie Libarkin, Ph. D.
Dina Lopez, Ph. D.
George Wood, Ph. D.

Statistical Analysis and Editing:
Sue Collins
Kizzi-Elmore Clark
John Aylesworth, Ph. D
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>4</td>
</tr>
<tr>
<td>List of Tables</td>
<td>6</td>
</tr>
<tr>
<td>List of Figures</td>
<td>7</td>
</tr>
<tr>
<td>Introduction</td>
<td>9</td>
</tr>
<tr>
<td>Purpose Statement</td>
<td>17</td>
</tr>
<tr>
<td>The Population and Sample</td>
<td>18</td>
</tr>
<tr>
<td>Definition of Terms and Variables</td>
<td>21</td>
</tr>
<tr>
<td>Data Collection and Analysis</td>
<td>24</td>
</tr>
<tr>
<td>Statistical Analysis of Scores</td>
<td>32</td>
</tr>
<tr>
<td>Results</td>
<td>34</td>
</tr>
<tr>
<td>Discussion of Results</td>
<td>53</td>
</tr>
<tr>
<td>Conclusions</td>
<td>61</td>
</tr>
<tr>
<td>References</td>
<td>66</td>
</tr>
<tr>
<td>Appendix A</td>
<td>69</td>
</tr>
<tr>
<td>Appendix B</td>
<td>78</td>
</tr>
<tr>
<td>Appendix C</td>
<td>80</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vocabulary Terms per Content Area on Ohio Graduation Test</td>
</tr>
<tr>
<td>2</td>
<td>Vocabulary Terms for Plate Tectonics Unit</td>
</tr>
<tr>
<td>3</td>
<td>Wilcox Data</td>
</tr>
</tbody>
</table>
List of Figures

Figures                      Page
1  Three-Tiered Approach to Classifying Vocabulary..........................15
2  Map Demonstrating Study Location in SE Ohio.................................19
3  Ohio Proficiency Test Scores for 2000-2001....................................19
4  GCI Example Question..............................................................26
5  OGT Example Question..............................................................27
6  Concept Mapping Instructional Slide.................................................28
7  Blank Student Worksheet Used for Concept Map Creation......................29
8  Mock Concept Map and ALA Reader Scoring Explanation........................30
9  Expert Concept Map.................................................................31
10 Student-Concept Map Example........................................................32
11 Full 23-Question Test Scores.........................................................37
    a. Pre-Test and Post Test Comparison............................................37
    b. Pre-Test and Post Test Comparison............................................37
    c. Pre-Test and Post-Test Comparison............................................38
    d. Pre-Test and Post Test Scoring Averages....................................38
12 15 Question Content Sub-Test Scores............................................40
    a. Pre-Test and Post Test Comparison............................................40
    b. Pre-Test and Post Test Comparison............................................41
    c. Pre-Test and Post-Test Comparison............................................41
    d. Pre-Test and Post Test Scoring Averages....................................42
13 15 Question Rasch GCI Sub-Test....................................................43
a. Pre-Test and Post Test Comparison ................................................. 43
b. Pre-Test and Post Test Comparison ................................................. 44
c. Pre-Test and Post-Test Comparison ................................................. 44
d. Pre-Test and Post Test Scoring Averages ........................................ 45

14 Low Scoring Concept Map ............................................................... 46

15 High Scoring Concept Map ............................................................. 47

16 Pre-Instructional Concept Map ......................................................... 48

17 Post-Instructional Concept Map ....................................................... 49

18 Concept Map Test Scores ............................................................... 51
   a. Pre-Test and Post Test Comparison ............................................. 51
   b. Pre-Test and Post Test Comparison ............................................. 51
   c. Pre-Test and Post-Test Comparison .......................................... 52
   d. Pre-Test and Post Test Scoring Averages .................................... 52
The effects of vocabulary intervention on ninth graders’ understanding of plate tectonics

Timothy James Sekula
Ohio University  Athens, Ohio
Email: ts189298@ohio.edu

“We use words to think; the more words we know, the finer is our understanding of the world” (Stahl, 1999)

Introduction

The Research Problem and Associated Studies

Progressive educational ideals such as those practiced by The Coalition of Essential Schools (CES)¹ provide students with the opportunities to investigate learning on higher levels of thinking while using one’s mind well. In a time of increased calls for single-test accountability from both No Child Left Behind (NCLB)² and the Ohio Graduation Tests (OGT)³, educators feel they are forced to move away from ideas set forth by progressive educational environments and focus their curriculum on attaining high standardized test scores. The ideas and goals set forth in the Ohio Academic

---

¹ http://www.essentialschools.org/pub/ces_docs/about/phil/10cps/10cps.html
³ http://www.ode.state.oh.us/proficiency/OGT/default.asp
Content Standards (OACS)\(^4\) for science provide a framework for what a scientifically proficient student should look like upon graduation from high school. However, the OACS fails to take into account the vast vocabulary content that is being included in the required standards and benchmarks.

A recent review (Table 1) of OGT preparation materials (Weikert, 2004) provides a look into the amount of content specific vocabulary which is expected to be taught by the end of the tenth grade year in all public high school science programs across the state of Ohio. The study materials under consideration stressed the importance of 629 “need to know” (Weikert, 2004) science vocabulary terms that any student preparing for the OGT should be familiar with. Table 1 easily demonstrates that secondary science programs are laden with immense amounts of terminology that can often be overwhelming for many students. For example, Earth Science related chapters list 86 vocabulary terms, most of which could prove to be new or rarely used language for many students.

---

\(^4\) [http://www.ode.state.oh.us/academic_content_standards/acssscience.asp](http://www.ode.state.oh.us/academic_content_standards/acssscience.asp)
Table 1. Graph Representing the Number of Vocabulary Terms per Content Area Possibly Appearing on OGT (Information Taken from OGT Study Guide Prepared by Orange Frazier Press; Weikert 2004)

Science courses are often considered to have their own “languages” (Lemke, 1989) and the most comprehensive introduction to this “language” appears in secondary science curriculums such as Biology, Chemistry, and the Earth Sciences. Groves (1995), mentions that knowledge of vocabulary is a fundamental key to understanding both the written and spoken language of science. Each chapter in a textbook or lecture in a classroom tells a story. For a student to truly grasp the “stories” being told in the science literature or in the classroom by instructors, they must be able to embrace the terms being taught and utilize them to make meaningful connections. Secondary science teachers often forego the notion of teaching how to read and comprehend terminology in a specific content area for the mere fact that such a heavy content load must be taught in order to adhere to state mandated standards. The vast amounts of information expected to be included in the curriculum often hinders a teacher’s ability to include important learning
tools such as vocabulary recognition, development, and understanding. In contrast, obtaining knowledge of word meanings and the ability to access that knowledge efficiently is important for comprehension (Chall, 1983,) and the depth of knowledge and understanding a student possesses in regards to the vocabulary might ultimately provide a direct correlation to their success on assessments.

Secondary science curriculums often measure student achievement by how well they acquire facts and information, but this information is often laden with technical vocabulary (Groves, 1995). Student achievement and understanding are difficult to assess when students cannot understand the content due to misunderstood terminology. It is important to find methods that to not only introduce students to these terms, but also actively engage them applying terms rather than just memorizing the meaning in the allotted time frame.

According to Marshall (1991), students may appear to understand words used in a science lesson, but often the interpretation of the meaning made by the student is not that which is expected by the teacher presenting the lesson. When students are faced with new pieces of material, they must take on the task of sorting between their previous background experiences and the meaning that accurately portrays the situation being presented in the classroom (Meagher, 1969). This conflict between conceptual understanding and new vocabulary can lead rise to the development of alternative, non-scientific conceptions. It is therefore of the utmost importance that teachers check and recheck their students’ understandings of essential vocabulary words (Marshall, 1991), so adequate comprehension of the subject matter can be ensured. An indirect or misdirected
emphasis of meaning can impede learning in science and may also lead to an incorrect or simplistic understanding of the nature of science (Groves, 1995).

A survey of secondary science texts conducted by Yager (1983) yielded some intriguing results in regards to vocabulary loading, the introduction of new language, in science curriculums and their associated textbooks. The study reviewed secondary science texts, discovering 2,173 and 17,130 new terms in a physical science and biology text, respectively. Although these counts may be somewhat elevated due to duplicate counting of words, this study suggests that significant mental energy must be expended by students to learn new vocabulary intertwined with science content. It is of interest to note that the amount of vocabulary utilized in the reviewed science texts is actually higher than that recommended for high school foreign language courses (Yager, 1983). This heavy use of terminology offers an obstacle to content comprehension as comprehension depends on a person already knowing 90-95% of the words in a text (Nagy & Scott, 2000). Overall, these studies suggest that misinterpreting vocabulary or the failure to obtain the meaning of vocabulary in a classroom can ultimately lead to misconception and inadequate comprehension of the subject matter.

In order to understand the types of vocabulary used in classrooms, Beck and McKeown (1985) developed a methodology for placing vocabulary into a three-tiered classification (Figure 1). The first tier refers to basic words which rarely require instruction of meaning (ex. talk, play, sad), whereas the second tier includes words that are considered “high-frequency” for mature, literate individuals that are found across a great range of domains (ex. ability, suggestion, transform). In general, the second tier words should be explicitly taught in classrooms simply because of their ability to expand
student vocabulary. The third tier refers to “low-frequency” words that are frequently limited to specific domains (e.g., genotype, rectilinear, isotope) and usually need to be pre-taught in order to help students obtain comprehension of domain specific content.

Secondary science curriculums are inundated with terms that fall within the third-tier realm. Beck and McKeown's (1985) work shadowed and expanded upon an earlier vocabulary framework presented by Dale (1965) known as the “Stages of word knowledge”. His classification system also placed many secondary science terms into categories that would require increased instruction for comprehension. As Stahl (1999) states, “We use words to think; the more words we know, the finer is our understanding of the world”. A student cannot be expected to grasp an understanding of the world without first being armed with the tools to achieve such a task. It is therefore obvious that integrated vocabulary tools are a necessary, yet often ignored part of earth-science curriculums in secondary science programs.
Vocabulary Selection
The Three-Tiered approach
(Bec & McKeown, 1985)

<table>
<thead>
<tr>
<th>First Tier:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The most basic words</td>
</tr>
<tr>
<td>• talk, play, sad</td>
</tr>
<tr>
<td>• Rarely require instruction in meanings in school</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Second Tier:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• High-frequency words for mature, literate individuals; found across a great range of domains</td>
</tr>
<tr>
<td>• ability, transform</td>
</tr>
<tr>
<td>• Should be taught explicitly because these words expand student vocabulary</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Third Tier:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Low frequency words; frequently limited to specific domains</td>
</tr>
<tr>
<td>• genotype, isotope</td>
</tr>
<tr>
<td>• Usually need to be pre-taught in order to help students understand the selection</td>
</tr>
</tbody>
</table>

Figure 1. Three-Tiered Approach to Classifying Vocabulary (From Beck & McKeown, 1985)

Vocabulary comprehension falls under one of the five pillars of reading comprehension set forth by the NCLB legislation. The “Reading First” program provides intensive reading instruction programs to students in grades K-3 and establishes the point that vocabulary terms are “…words that a student must know in order to communicate effectively” (Report of the National Reading Panel, 2000)⁵. This program establishes the notion that creation of a quality and proficient reader can be established in early academic years but abandoned after only the third grade at which point specific academic content is focused upon rather than literacy methodologies. It is in the later academic years (secondary schooling) where students encounter some of the highest levels of vocabulary

inundation, yet receive minimal instruction on how to integrate it into their learning. The contradiction here is that vocabulary development is abandoned just at the time when it would have the highest payoff for students. Understanding that “Knowledge of vocabulary is a key to understanding both spoken and written language” (Johnson & Piercey, 1984), it becomes essential for secondary science programs to recognize the importance of not only teaching vocabulary but also integrating it throughout the entire curriculum as means of providing the necessary tools for meaningful connections to be established. Vocabulary instruction in secondary science classrooms may be a key to improving student comprehension.

The goal of vocabulary instruction is to establish deep word knowledge that allows students to integrate the meanings with prior knowledge while providing multiple word exposures (Nagy, 1988). Vocabulary interventions in the classroom may provide students and teachers the opportunity to investigate scientific concepts while using vocabulary comprehension as a backdrop for understanding. A vocabulary intervention can focus on many different aspects of the learning process and include a variety of variables such as (Bryant, 2003):

1.) Word Selection Procedures
2.) Differing Instructional Materials
3.) Instructional Design and Procedure
4.) Duration of Instruction
5.) Measures of Vocabulary Learning (Word Knowledge Level)

Each and any of these variables may be utilized before, during, or after instruction in an effort to deepen as well as gauge understanding of the content under consideration.
Curtis and Longo (2001) warn those engaging in vocabulary intervention to be careful to choose words with “high utility”, words that will be encountered again and again, and also to chose a relatively low frequency of terms. The necessity of focusing on a limited number of terms presents a problem to secondary science teachers who are often forced to overwhelm their students with new, narrow, or content specific vocabulary embedded in the state standards.

Two studies conducted by Sahiner & Schneps (1977 and 1997) investigated levels of science aptitude amongst graduating students from Harvard University and MIT using elementary level content and questions. It was found that students at Harvard were unable to explain the cause of seasons on Earth and MIT students unable to explain simple elementary biological principles. These studies identify a problem that exists not only in the primary educational years, but also in the post-secondary years. The methodology being utilized to teach the overwhelming breadth of content that exists in the sciences is deficient in attaining its purpose. The teachers and instructors are failing to provide our students with the necessary background to explain everyday, common scientific principles.

Purpose Statement

This study investigates the effects of vocabulary intervention techniques on student comprehension and assessment rates (pre- and post-test scores) for a high school earth-science unit on plate tectonics. Students were randomly assigned to three groups: 1) a control group, where students were exposed to lectures on plate tectonics; 2) experimental group 1, where students were exposed to lecture, inquiry activities and
multiple vocabulary intervention tools; and 3) experimental group 2, where students were exposed to lectures and inquiry activities without language intervention. The main goal of the study was to identify if the vocabulary interventions produced an increase in student comprehension and connections of plate tectonics in comparison to the groups not exposed to the language interventions. Comprehension was evaluated through comparison of quantitative standardized test and concept map scores. This study investigated whether vocabulary intervention yielded higher comprehension rates and assessment scores for the students receiving vocabulary intervention versus those who do not receive it throughout the unit. The focus of this study is on literacy and effects of vocabulary intervention, but did not only center on vocabulary related to reading. This study also examined the importance of integrating vocabulary throughout the unit using a variety of learning methods that address reading, writing, speaking, and listening skills.

The Population and Sample

The population under study consisted of ninth grade, earth science students ranging in age from 14-16 years old enrolled in a public high school in rural southeastern Ohio (Fig. 2). This school has a population of approximately 420 students of which approximately 92.6% are considered to be white or non-Hispanic, 3.7% black, non-Hispanic, and 3.4% multi-racial.
The high school and surrounding geographical setting is generally characterized by below average family income (52.9% economically disadvantaged), above average poverty levels and number of identified students (those who qualify for special education services 15.9%), and standardized test scores that annually fall below the state of Ohio calculated averages\(^1\). The Ohio Proficiency Test was administered in 2001 to seniors who attended school in the study area. Their scores fell well below the state average of 60% (Figure 3) demonstrating the characteristic low scores associated with the area. In 2005 the Ohio Graduation Test (OGT) was administered to tenth grade students from this district and yielded scores in math, reading, writing, and science below the state averages.

![Figure 2. Map of Study Location in South-Eastern Ohio](image)

![Figure 3. Chart Demonstrating Below Average Test Scores on Ohio Proficiency Tests for Students in the Study Area for the 2000-2001 Academic Year.](image)
for the 2004-2005 academic school year. The relatively high amount of students in the study facing difficulties either from learning disabilities or financial inequalities creates a unique population that may not represent an “average” population of students. This particular population was chosen because of the convenience and availability of the students in relation to the researcher as well as the possibilities of applying the findings of the study to benefit the study population.

The students participating in the study were currently enrolled in a Science 9 (Freshman Physical Science) course that covers earth-science content. The population of n = 54 students were pulled from their normal daily teacher and randomly assigned to one of three groups (Control, Experimental 1, Experimental 2) being utilized in the study creating three groups of n = 18 students each. Random assignment of students was conducted through the Microsoft EXCEL random number generation function. Each of the three groups of students were then randomly assigned to one of the three instructors. It must be noted that in order to maintain the characteristics of a True Experiment (Creswell, 2003, p. 170) all students had to be randomly assigned to groups and therefore some of the students ended up with their normal, everyday teacher. Also, of the three male teachers participating in the study, only two of the instructors were currently teaching the students being used for the research. Students remained with the assigned teacher for a total of six instructional periods over the course of a normal school week. The district allowing the study participates in a block-scheduling system which allows for 65-80 minute class periods on a daily basis. This particular unit was designed to have students in the classroom for a total of about 450 minutes of instructional time.

Ohio Department of Education, 2005-2006
Definition of Terms and Variables

Students participating in this study were exposed to lectures related to the plate tectonics curriculum. Students were exposed to all the required content as specified by the OACS. Instructors were provided with all the necessary materials to teach the course including power-point slide shows, vocabulary intervention activities, worksheets, hands-on activities, and pre- and post-test assessments. The structure of the lessons allowed for minimal influence of teaching style and characteristics of the instructor. Every student participating in the study was asked to complete pre- and post-instructional tests and concept maps.

The researcher in this study created all curriculum materials and distributed them to the proper instructor participating in the study (Appendices A, B, and C) A general question and answer session was held before the study took place to become familiar with the lessons and to distribute materials. The instructors were not informed as to the purpose of the study, and were asked to teach only according to the prescribed lesson plans. At the closure of each lesson, instructors provided feedback on the quality of the materials, apparent effectiveness of lesson delivery, and day to day classroom issues such as attendance and work completion rates.

The instructors in this study were all male ranging in age from 21 years old to 45 years old. Two of the teachers received their teaching credentials from Ohio University in Athens, Ohio while the third received out of state accreditation.

- Control Group
  - This population of 18 students served as the control for this study.

Students were exposed to all the required content as specified by the
OACS, but did not receive vocabulary interventions or actively engage in learning activities. Students did receive any treatment other than general, lecture based instruction as part of a normal curriculum. These students participated in the pre and post-test assessments.

• Control Group Teacher
  o The instructor for the control group was a veteran teacher with nine years of high school teaching experience. The teaching qualifications for this instructor include a B.S. of Environmental Studies as well as a Masters of Teaching and Learning. In addition to these qualifications, this instructor has experience teaching Science 9, Science 10, Biology, and Physics in the district being utilized in the study. This instructor was chosen for assignment to this group to exclude any possibility of years of experience producing positive effects in the experimental groups.

• Experimental Group 1
  o This population of 18 students served as one of the two experimental designs for this experiment. Students were exposed to six brief lectures of 20 minutes each, received multiple language interventions throughout the unit, and participated in inquiry-based activities. The language interventions included such activities as word walls, elicitation of background knowledge by the instructor, word sorts, and take home worksheet activities.
• Experimental Group 1 Teacher
  o The instructor for experimental group one was a first year teacher, new to
    the field of science teaching although had one quarter of student teaching
    experience. The teaching qualifications for the instructor include a BS of
    Life Science as well as a BS in Wildlife Biology in addition to being a
    teacher of Science 9, Science 10, and Environmental Science in the district
    being used in the study.

• Experimental Group 2
  o This group of 18 students served as the second of two experimental design
    groups for the experiment. Students received brief lectures of 20 minutes
    each and participated in inquiry based activities just as experimental group
    # 1 experienced. These students were not exposed to any vocabulary
    intervention activities as experienced by the students in experimental
    group # 1.

• Experimental Group 2 Teacher
  o The instructor for experimental group two was a student teacher working
    on his in-service training experience, pursuing a BS in Education with an
    emphasis in Integrated Science. This instructor was chosen for
    assignment to this group because of lack of experience in a classroom
    setting. It was felt by the researcher that this instructor would be unable to
    complete all of the required activities set forth in experimental group one,
    but would be able to successfully complete those laid out in experimental
    group two.
Data Collection and Analysis

The collection of pre and post-instruction quantitative and qualitative data was utilized in this study in order to obtain an applicable statistical interpretation of the results. The use of pre-instructional and post-instructional testing provided the opportunity to gauge student performance in each study group (Experimental 1, Experimental 2, and the Control Group) and to determine if sufficient gains or losses occurred as a direct result of the experienced instructional methodology. The data collection format was divided into two distinct assessments that consisted of a 23 question standardized test (Appendix A) and the use of concept mapping. Each assessment format was scored quantitatively using standard approaches.

The standardized test format was chosen as a method to assess student achievement because it was easily quantifiable and could be utilized in comparisons with other studies. The students being utilized in the study were familiar with the standardized testing format and therefore required minimal instruction for completion of the assessment. The multiple-choice, standardized test created for the study consisted of a combination of 23 questions gathered from recent Ohio Graduation Tests (OGT) and from the Geological Concept Inventory (GCI)\(^7\). The GCI questions utilized were selected from a validated pool of 73 questions developed for use in earth science classrooms that include test items covering general physical geology concepts as well as underlying fundamental ideas in physics and chemistry (Libarkin et al., 2005; Libarkin and Anderson, 2005a,b). Each GCI (Figure 4) question has been through a validation process known as Rasch testing procedures for validation. Rasch analysis has been used to assist

\(^7\) [http://newton.bhsu.edu/eps/gci.html#A](http://newton.bhsu.edu/eps/gci.html#A) (Libarkin et al., 2005; Libarkin and Anderson, 2005a,b).
in the development of 15-item GCI sub-tests from the bank of 73 GCI questions with the idea of addressing scale linearity. Pre- and post-instruction standardized 15 item GCI sub-test score were scaled using:

\[ S_{GCI} = 16.76 + 4.30R_{GCI} + 0.115(R_{GCI}-7.5)^3 + 0.042(R_{GCI}-7.5)^4 - 0.0017 (R_{GCI}-7.5)^5 \]

where \( S_{GCI} \) is the scaled score on a 0-100% scale and \( R_{GCI} \) is the raw score on a 15-item GCI sub-test. The 23-question test created for this study included 15 GCI questions randomly distributed throughout the test that were selected based upon requirements set forth by Libarkin and Anderson (2005). The scores for the 15 question GCI sub-test were analyzed as scaled scores rather than raw score percentages which allows for the idea of scale linearity to be included.

The 15 question Content Subtest was created from the pool of 23 GCI/OGT questions in an effort to assess student gains on plate tectonic related questions only. The GCI questions utilized offered some variability as to topic coverage because certain questions must be utilized on the test to allow for scaling from raw to scaled scores. This scaling allows for comparison with other GCI sub-tests used by other researchers. The 15 questions selected for the Content Subtest were plate tectonics based questions that had direct correlation with the material being delivered in the three different study classrooms. The scores for this 15 question Content Subtest were analyzed as raw percentage scores for both the pre- and post-tests.

In addition to the 20 GCI questions chosen, 3 OGT questions (Figure 5) were selected for use on the standardized test. These questions were chosen from previous standardized tests provided to students by the state of Ohio over the past two years. The
three questions addressed specific plate tectonics content that was directly related to the unit being taught. The scoring of these questions were included in the analysis of the Full 23 question Test, but were not selected for individual analysis.

*The map below shows the position of the Earth’s continents and oceans today. The gray areas represent land, and the white represents water. Which of the following best explains why the ocean basins look the way they do?

(A) Meteor impacts caused the ocean basins to form this way
(B) Ocean basins form as continents move
(C) The ocean basins formed in cracks that were created as the whole Earth cooled after its formation
(D) The ocean basins formed in cracks that were created as the whole Earth heated after its formation

*Figure 4. Example of Question from the Geoscience Concept Inventory*
The second assessment tool utilized in this study consisted of the use of pre and post instructional concept mapping activity that all students were asked to complete. Each student was provided with the same pre-determined list of 24 vocabulary terms (see below) consistent with the content being studied throughout the unit as well as a focus topic for the concept map: “What Processes Shape the Earth?”

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Continental Drift</td>
<td>Ridge</td>
<td>Lava</td>
</tr>
<tr>
<td>Convergent Boundary</td>
<td>Plate Tectonics</td>
<td>Continental Plate</td>
</tr>
<tr>
<td>Divergent Boundary</td>
<td>Igneous</td>
<td>Hot Spot</td>
</tr>
<tr>
<td>Earthquake</td>
<td>Volcano</td>
<td>Magma</td>
</tr>
<tr>
<td>Fault</td>
<td>Tsunami</td>
<td>Oceanic Plate</td>
</tr>
<tr>
<td>Rift</td>
<td>Gravity</td>
<td>Trench</td>
</tr>
<tr>
<td>Sea-Floor Spreading</td>
<td>Asthenosphere</td>
<td>Convection</td>
</tr>
<tr>
<td>Subduction</td>
<td>Density</td>
<td>Tectonic Plate</td>
</tr>
</tbody>
</table>

Concept maps were chosen as an assessment tool for this study because they provide a methodology to assess organization and representation of conceptual ideas associated
with the unit of study. Each group participating in the study was provided with instruction (Figure 6) on how to construct concept maps during the first day of the study.

![Components of a Concept Map](http://cmap.ihmc.us/Publications/ResearchPapers/TheoryCmaps/Fig5CmapSeasons-large.png)

**Figure 6. Example Used to Instruct Students on how to Construct Concept Maps.**

Fifty-minutes of class time was devoted to the instruction and creation of the pre concept maps on the first day while 35 minutes of class time was allotted for creation of the post-concept maps on the final day of instruction. The fifteen minute difference can be accounted for due to the necessary instructional time for teaching concept map creation. Each student was provided with a pre-fabricated worksheet (figure 7) upon which to create their concept maps. The worksheet provided the students with the vocabulary list of 24 terms, the focus topic, and a blank area in which to build and construct the map. Students are provided with the “concept words” and “focus topic” and are asked to demonstrate how the words connect with each other by providing “linking words” and “connecting lines”. The focus topic for this unit was an inquiry that posed the question “What processes shape the Earth?”. Students provided linking terms that helped create a written-like approach to demonstrating their understanding of the topics.
It is difficult to quantify assessment pieces such as concept maps due to the inherent objectiveness and variability of those grading them. The creation of CMAP Tools developed at the “Institute for Human and Machine Cognition” and ALA Reader developed by Dr. Roy Clariana at Penn State University has allowed for a completely subjective value to be assigned to this type of assessment with the application of minimal effort. The concept maps created by the students (Figure 10) were done by hand due to the lack of available computer technology in the study area but were then converted into electronic CMAP files (Figure 8) so a quantitative analysis could be performed.

CMAP Tools allows for the created files to be exported as text (.txt) proposition files that arrange the contents of the map into a text based format that can be assessed using ALA Reader (Clariana, 2004). ALA Reader uses an algorithm called “painaa text summary utility (Pathfinder analysis of inferential aggregated associations)” that converts

---

8 http://cmap.ihmc.us/
9 http://www.personal.psu.edu/faculty/r/b/rbc4/score.htm
sentences in a written text into co-occurrence vectors and then converts each co-occurrence vector into its resulting propositions (Figure 8). The raw score created through ALA Reader is an agreement score with the expert map. As seen in figure 7, if term A links to term C then 1 point is given to the student. If a student links term A to term D, it does not agree with the expert map and therefore 0 points are given to the student score. Similar studies using such quantitative assessment tools can be seen in the work of Clariana, R.B., & Koul, R (2004).

Figure 8.  Mock Concept Map (Left), the Propositions-as-Text File of this Map (Center) Generated Using the Export-Cmap-As feature, and the Proximity Array (Right) Generated by ALA-Reader from this Text File (Clariana, 2005)

Student (.txt) files are scored in ALA Reader against an “expert” file created from an original “expert” concept map (Figure 9). The expert map was created by the three teachers (Experimental # 1, Experimental # 2, and Control) participating in the study to ensure heterogeneity and fairness in the scoring process. Each teacher created a map from the provided terms before the study took place and then compared their maps for similarities and differences. The expert map created for this study was a compilation of all of the similarities, as judged by the experts, that existed between the three different teachers maps. The comparison between the student .txt files and the expert .txt file is a “rudimentary”
scoring system that provides a percent agreement with the expert file. The scoring of pre- and post instruction concept maps allows for instruction-related changes to be investigated amongst the students participating in the study.

Figure 9. Expert Concept Map Utilized for Comparison Against Student Compiled Concept Maps and Scoring Purposes using ALA Reader.
Figure 10. Student Concept Map Example that is Compared to the Expert Map for Agreement Scoring

Statistical Analysis of Scores

The pre- and post assessment scores for both the concept mapping and 23-question standardized test were analyzed using The Wilcoxon Signed Rank Test (McClave, 2000) to check for significant gains amongst the three instructional groups participating in the study. The Wilcoxon Signed Rank Test, also known as the Wilcoxon Matched Pairs Test is a non-parametric test used to test the median difference in paired data (Shaw, 2000). This test is the non-parametric equivalent of the paired t-test. Non-parametric statistics were necessary in this study because of the small sample sizes of the three study groups (n = 18).

The null hypothesis, $H_0$, for the Wilcoxon Signed Rank Test states that the scoring distributions for the pre and post-tests are identical. If the null hypothesis is true,
the sum of the ranks of the positive differences should be about the same as the sum of the ranks of the negative differences and the values calculated can be compared with critical values established in the statistical tables (Shaw, p. 761, 2000). In order to reject the null hypothesis, the calculated value must be lower than the critical value. If the null hypothesis (H₀) is rejected, the data will suggest that the pre- and post-test score distributions are different with respect to each other.

The small population sizes available for study created the opportunity for non-parametric statistics to be used to analyze the populations and determine if significant gains occurred between pre- and post-instruction. The Wilcoxon Signed Rank Test was used to evaluate the following assessment items for each study group:

- Pre vs. Post Concept Map Percent Agreement
- Pre vs. Post Full GCI/OGT Test (23 Questions)
- Pre vs. Post 15 Question GCI RASCH Subtest
- Pre vs. Post 15 Question Content Subtest

The concept mapping assessment yielded results that required further statistical analysis to investigate the significance of the scoring outcomes. The Kruskal-Wallis H-Test was utilized to compare the pre-test and the post-test raw scores to determine if the probability distributions were the same for all study groups (Shaw, 2000). The null hypothesis (H₀) for the Kruskal-Wallis H-Test states that the probability distributions of the study populations are the same for all three study groups (experimental group # 1, experimental group # 2, and the control group). The calculated H-statistic must be greater than the critical value listed in the table in order to reject the null hypothesis (H₀). The
purpose for utilizing this test was to determine if the study populations from each group had the same probability distributions on the pre-test but different on the post-test.

Results

The random design utilized in this study created three distinct but equal study populations (n = 18, n = 18, and n = 18) from the available sample size (n = 54). Students were removed from statistical analysis if they missed either a pre- or post-test due to absences. The resulting population sizes after removal of students missing an assessment included the following: Experimental Group # 1 (n = 15), Experimental Group # 2 (n = 18), and the Control Group (n = 16). In addition to removal of students who missed either a pre- or post-test, the elimination of students who achieved an absence rate $\geq 50\%$ were eliminated from some statistical analysis to eliminate possible outliers from the data sets. The final populations for this study included experimental group # 1 (n = 11), experimental group # 2 (n = 16), and the control group (n = 15).

The full 23 question GCI/OGT test was characterized by relatively low overall scores (Figure 11) on both the pre- and post-test assessments but had a wide range of scores falling between 4.3% and 65.2%. The high-end score of 65.2% was achieved by only one student on the post-test whereas the low-end score of 4.3% was achieved by one student on the post-test as well. Nearly 25% of the students in the population (n=42) received a score below 20%. The post-test scores showed a range from 4.3% to 65.2% each having only one student obtain each score. Although the scores appear to be relatively low, it must be noted that scoring gains took place amongst all three study groups in the population sample. The scores show that the students had some knowledge
of plate tectonics content before the instructional unit and increased their content understanding overall after the instructional period.

Experimental group #1 experienced scoring gains amongst 8 out of 11 students participating in the assessment. The average pre-test score was a 28.46% and the average post-test score was a 38.74 providing an average scoring increase of 10.2. This was overwhelmingly the largest gain posted by any group in the study. Although experimental group #1 did not have the lowest pre-test score (27.83% by the control group), the students managed to obtain the highest post-test score average amongst any group in the study (38.74%) for the full 23 question test. One student in the group posted a pre-test score of 21.7% and a post-test score of 52.1% measuring a scoring gain of 30.4%. The ranges of scores for the group on the post-test went from 21.7% to 60.8%. The significance of the scoring gains can be seen through the use of the Wilcoxon test demonstrating significant gains at $\alpha = .10$ and .05 showing increased understanding of the instructional content material on plate tectonics.

Experimental group #2 demonstrated minimal scoring increases between the pre- and post-test scores for the full 23-question test although 8 out of the 16 students (50%) showed some scoring increase between pre- and post-instruction. The average pre-test score was 32.61 and the average post-test score was a 35.60% for a 2.9% average gain in score. The population of 16 students showed some division amongst it scores where 8 students maintained or decreased in score whereas 8 students showed some scoring gain between pre- and post-test scores. The students in this group posted the highest pre-test score at 32.61% but maintained the middle value score amongst the three study groups for the post test-assessment where experimental group #1 posted the highest score at
38.74%. Students in this group found a pre-test scoring range from 17.3% to 60.8% with only one student posting the high-end score. On the post-test assessment scoring ranges decreased to fall in the range of 4.3% to 56.5% showing that high end pre-test scores actually showed some scoring loss from pre- to post-instruction. The use of Wilcoxon statistical analysis demonstrated no significant gains for this population of students.

The control group measured an average pre-test score of 27.83% and an average post-test score of 34.78% showing an average increase of 6.9%. The students of this population maintained the lowest average pre-test score of 27.83% and found a range of scores from 8.6% to 47.8%. This group of students showed the lowest pre- and post-test scores amongst all groups in the study although 9 of the 15 (60%) students showed a scoring increase between pre- and post-instruction. The post-test scores showed a high end of 60.8% and a low-end of 13.0% showing overall increases in scoring. The use of Wilcoxon statistical analysis shows that significant gains occurred for this population at the $\alpha = .10$ and $.05$ levels showing increased understanding of the content.
11a. Represent Individual Student Scores and those Students Falling Above the Trend-Line Experienced Scoring Increases While those Below Experienced Scoring Decreases.

11b. Represent Individual Student Scores and those Students Falling Above the Trend-Line Experienced Scoring Increases While those Below Experienced Scoring Decreases.
11c. Represent Individual Student Scores and those Students Falling Above the Trend-Line Experienced Scoring Increases While those Below Experienced Scoring Decreases.

11d. Represents the Average Scores Achieved by Students in each Study Group for both the Pre- and Post-Tests.
The 15 question content sub-test represented the highest overall average scores (Figure 12) for this study, with a low average score of 30.83% on the pre-test and a high score of 47.87% on the post-test for the overall population. The 15 questions utilized were expected to demonstrate higher rates of gain amongst students in the study (n = 42) because they were all related to content delivered throughout the instructional unit. Student pre-test scores posed a low-end score of 6.67% (n=4) and high-end score of 66.67% (n=2). The post-test scores ranged from a low-end score of 0.0% (n=1) to a high-end score of 80.0% (n=3). The largest scoring gain on the assessment from pre- to post-test was 40% posted by two students in the study. Of the 42 student scores analyzed in this sub-test, 24 demonstrated an increase from pre- to post-test scoring. The largest average gain for any group in this study was 15.1%, and the lowest was a 6.6% increase.

Experimental group # 1 (n=11) showed scoring gains on the 15 question content sub-test from 7 out of its 11 students (63%) and had 0 students show scoring decreases. Students in this group maintained an average scoring increase from pre- to post-test of 15.1% but the highest average scoring increase was 26.6% achieved by two students in the population. Students in this population demonstrated post-test scores ranging from 20% to 73.3%. The low-end score of 20% was achieved by only 1 student and this student recorded no change in score from pre to post test scoring. Analysis of the scores using the Wilcoxon test showed levels of significance to the $\alpha = 0.01$ or 99% confidence level.

Experimental Group # 2 (n = 16) experienced scoring increases from 7 of its 16 students (43%) and showed an average overall scoring gain of 6.6%. Student scores showed an average pre-test score of 32.59% and an average post-test score of 39.25% but
had 3 students achieve the high-end score of 80%. The 3 students who obtained the highest post-test scores also had some of the highest pre-test scores therefore the average scoring gain for the population is relatively small even though high scores were achieved. The scoring increase of 6.6% was shown to be statistically significant at the $\alpha = 0.05$ level or 95% confidence interval according to the Wilcoxon test.

The control group (n = 15) students showed a scoring increase in 10 of its 15 students (66.6%) with the highest post-test score being a 73.3% from two students. The low-end pre-test score was a 20%. The average scoring increase amongst students in the population was 7.97% where the average pre-test score was 30.83% and the average post-test score was 38.8%. The largest scoring increase was 26.67% posted by three students in the study population. The use of the Wilcoxon statistical test shows the scoring increases to be significant at the $\alpha = 0.10$ level or 90% confidence interval.

12a. Represent Individual Student Scores and those Students Falling Above the Trend-Line Experienced Scoring Increases while those Below Experienced Scoring Decreases
12b. Represent Individual Student Scores and those Students Falling Above the Trend-Line Experienced Scoring Increases while those Below Experienced Scoring Decreases

12c. Represent Individual Student Scores and those Students Falling Above the Trend-Line Experienced Scoring Increases while those Below Experienced Scoring Decreases
The 15 question Rasch GCI sub-test (Figure 13) reported here are scaled as described and are therefore comparable to other GCI sub-test data. Of the 42 students participating in the assessment process, 26 demonstrated scoring increases from pre- to post-test assessment. The lowest average pre-test score was a 26.55% and the highest average pre-test score was a 33.64%. The study groups experienced average scoring increases from 6.6% to 15.1% for the 42 students participating in the study. Individual student pre-test scores ranged from 0% to 55.8% and post-test scores range from 11.3% to 61.0% of which each score was registered only once.

Experimental group # 1 was characterized by 8 out its 11 students experiencing gains in scoring from pre-test to post-test. Four students experienced 0% gains while 1 student experienced a gain of 35.53% on this assessment. Average pre-test scores for this group were 32.7% while the post-test score was a 47.87% showing a 15.1% scoring gain. This gain was nearly double of any other group in the study for % gains on the 15 question Rasch GCI sub-test. The lowest post-test score for these students was a 27.46% while the highest was a 55.8%. One student in the study group showed a 33% increase from a 0% pre-score to a 33% post-test score. Wilcoxon statistical analysis showed that
the scoring gains experienced by students can be seen as significant to the $\alpha = 0.01$ level or 99% confidence interval.

Experimental group # 2 experienced scoring gains from 9 of its 16 students with an average scoring increase from pre- to post-test of 4.4%. Students recorded an average pre-test score of 33.64% and an average post-test score of 38.13%. Although 9 students experienced scoring gains, four students also experienced scoring decreases ranging from -10 to -16.1%. The largest scoring gain for this group was a 17.96% increase from pre- to post-test. The use of the Wilcoxon statistical test showed that the gains experienced by this group were not significant in nature.

The control group for this study consisted of n=15 students of which 9 experienced scoring gains on the 15 question Rasch GCI assessment. The average pre-test score was a 29.5% and the average post-test score was a 37.47 showing an average scoring gain of 8.32%. Students in the group showed scoring increases of 20.16% and 27.76% but were not the highest gains in the study. 5 students showed gains of 15% or greater demonstrating a deeper understanding of the delivered content for the unit on plate tectonics. The Wilcoxon non-parametric test shows that the scoring increases can be viewed as significant to the $\alpha = 0.01$ level or 99% confidence interval.

![Experimental Group # 1](image)

13a. Represent Individual Student Scores and those Students Falling above the Trend-Line Experienced Scoring Increases while those Below Experienced Scoring Decreases
Experimental Group # 2

15 Question GCI RASCH

13b. Represent Individual Student Scores and those Students Falling above the Trend-Line Experienced Scoring Increases while those Below Experienced Scoring Decreases

Control Group

15 Question GCI RASCH

13c. Represent Individual Student Scores and those Students Falling above the Trend-Line Experienced Scoring Increases while those Below Experienced Scoring Decreases
13d Represents the Average Scores Achieved by Students in each Study Group for both the Pre- and Post-Test.

Concept maps created by the study population (n=42) prior to instruction indicate that students, although familiar with language relevant to plate tectonics, were unfamiliar with relationships between the concepts. 14 students (33.3%) created concept maps that received a pre-test score of 25% agreement with the expert map or less and the highest pre-instruction score was only a 58%. In general, the maps were dominated by extreme variance in both structure and inclusion of concept words. Student concept maps varied from the inclusion of only three terms (Figure 9) to the use of all 24 terms (Figure 10) with all others falling in between these two extremes. The structure in which the students presented their concept maps did not play a major role in the scoring process as only links, not hierarchal structure, were scored. The higher scoring maps did represent a number of key concepts necessary for understanding the unit on plate tectonics but
overall the students to identify with the expert map. It is believed that the student’s lack of deep understanding of the vocabulary in the unit inhibited their ability to perform well on the pre-test assessment.

Figure 14. Example of a Concept Map Using Only Three Terms. This Map Received an 8.3% Agreement with the Expert Map.
After instruction, the majority of students (62%) experienced no change or a decrease in their concept mapping scores. However, of the 16 students that demonstrated a gain in their concept mapping scores from pre- to post-test, 56% (n=9) were in experimental group # 1 while 25% fell in experimental group # 2 and 19% in the control group. These students exhibited an increased ability to make connections between concepts presented in the unit on plate tectonics. This can be seen when comparing a students pre- and post-test assessments (Figure 14) from experimental group # 1. Figures 11 and 12 demonstrate the pre-test map with a score of 33% agreement and the post-test map with 58% agreement.
Figure 16. A Pre-Instructional Concept Map that Received a 33% Agreement with the Expert Map for a Student in Experimental Group # 1.
Experimental group #1 (n=11) experienced scoring gains by 81% of its students. The average pre-test score was a 29.55% and the average post-test score was a 40.15% providing an average overall gain in score of 10.6%. The use of the Wilcoxon statistical analysis shows that gains experienced by this group of students can be considered significant at the $\alpha = 0.02$ level or 98% confidence interval. In order to justify these results, the Kruskal-Wallace H-Test was utilized. This analysis for the pre-test scores accepted the null hypothesis ($H_0$) stating the probability distributions for all three study groups were identical. The same analysis was utilized on the post-test scores for all three study groups and it rejected the null hypothesis ($H_0$) identifying that there is sufficient
evidence to suggest that the populations have different probability distributions for the post-test assessment. The populations were statistically the same on the pre-test therefore demonstrating the significance of the gains on the post-test. These non-parametric statistical tests allow for the idea to arise that experimental group # 1 experienced significant scoring gains as a study group.

Experimental group # 2 (n=16) experienced scoring gains by only 25% of its students while the remaining 75% either maintained their score or decreased from pre- to post-test. The average pre-test score was a 23.96 and the average post-test score was a 21.35%, indicating an average overall decrease in score of -2.6%. The use of the Wilcoxon statistical analysis test showed no significant scoring gains or losses on the concept mapping assessment at any level of significance.

The control group (n=15) demonstrated scoring gains by only 20% of its students and is characterized by the fewest gains amongst all groups in the study. With an average pre-test score of 30.00 and average post-test score of 26.66, it can be seen that the -3.33% average score decrease was the largest of the study groups. The use of the Wilcoxon statistical analysis test showed no significant scoring gains or losses on the concept mapping assessment at any level of significance.
18a. Represent Individual Student Scores and those Students Falling Above the Trend-Line Experienced Scoring Increases While Those Below Experienced Scoring Decreases.

18b. Represent Individual Student Scores and those Students Falling Above the Trend-Line Experienced Scoring Increases While Those Below Experienced Scoring Decreases.
18c. Represent Individual Student Scores and those Students Falling Above the Trend-Line Experienced Scoring Increases While Those Below Experienced Scoring Decreases.

18d. Represents the Average Scores Achieved by Students in each Study Group for Both The Pre- and Post-Tests.
Statistical Analysis: Wilcoxon Results

### Experimental Group # 1

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Average Pre-Test Score (%)</th>
<th>Average Post-Test Score (%)</th>
<th>Wilcox T-Statistic</th>
<th>$\alpha = .10$</th>
<th>$\alpha = .05$</th>
<th>$\alpha = .02$</th>
<th>$\alpha = .01$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMAP</td>
<td>29.55</td>
<td>40.15</td>
<td>0</td>
<td>11</td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>15 Q</td>
<td>32.72</td>
<td>47.87</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Content</td>
<td>28.46</td>
<td>38.74</td>
<td>7.5</td>
<td>11</td>
<td>8</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Full Test</td>
<td>26.55</td>
<td>39.1</td>
<td>0</td>
<td>11</td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

### Experimental Group # 2

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Average Pre-Test Score (%)</th>
<th>Average Post-Test Score (%)</th>
<th>Wilcox T-Statistic</th>
<th>$\alpha = .10$</th>
<th>$\alpha = .05$</th>
<th>$\alpha = .02$</th>
<th>$\alpha = .01$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMAP</td>
<td>23.96</td>
<td>21.35</td>
<td>30.5</td>
<td>17</td>
<td>14</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>15 Q</td>
<td>36.67</td>
<td>44.16</td>
<td>5.5</td>
<td>8</td>
<td>6</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Content</td>
<td>32.61</td>
<td>35.6</td>
<td>36.5</td>
<td>26</td>
<td>21</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Full Test</td>
<td>33.04</td>
<td>38.13</td>
<td>21.5</td>
<td>11</td>
<td>8</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

### Control Group

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Average Pre-Test Score (%)</th>
<th>Average Post-Test Score (%)</th>
<th>Wilcox T-Statistic</th>
<th>$\alpha = .10$</th>
<th>$\alpha = .05$</th>
<th>$\alpha = .02$</th>
<th>$\alpha = .01$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMAP</td>
<td>30</td>
<td>26.66</td>
<td>20</td>
<td>11</td>
<td>8</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>15 Q</td>
<td>32.88</td>
<td>41.33</td>
<td>25</td>
<td>26</td>
<td>21</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Content</td>
<td>27.83</td>
<td>34.78</td>
<td>8.5</td>
<td>14</td>
<td>11</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Full Test</td>
<td>29.15</td>
<td>37.47</td>
<td>0</td>
<td>11</td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3. Wilcoxon Matched Pairs Statistical Analysis for Each Study Group. The Cells Highlighted in Red Demonstrate Regions where Significant Gains were Found to Exist. The Red Cells Represent Areas where the Calculated T-Stat was Lower Than the Critical Value Therefore Rejecting the Null Hypothesis.

Discussion of Results

The results discussed below contain a look at the results with and without student absences. It was decided that students having an absence rate $\geq 50\%$ would be eliminated from the final reported results as these points appeared to be outliers in comparison to those students falling below the 50% absence rate. In addition to the absent students
being eliminated from the final reporting, any student missing either the pre- or post-tests had to be eliminated from the study because valid analysis could not be performed upon these students. As was initially expected, experimental group # 1 demonstrated the largest gains on both the standardized test and concept mapping assessments however the control group posted some significant gains as well.

Concept maps allow for examination of valid understandings and misconceptions that students may hold (Zeilik, 1998). The maps created by the students in this study provided an excellent tool for comparison of the effectiveness of instruction provided to the different study groups. The removal of students having an absence rate $\geq 50\%$ provides some valuable insight into the effect of vocabulary interventions on students understanding of plate tectonics in relation to the concept mapping assessment. Using the Wilcoxon Signed Rank Test to assess student pre- and post-test scores, it becomes evident that the vocabulary interventions implemented in experimental group # 1 provided an effective tool for addressing and teaching plate tectonics related concepts. The rejection of the null hypothesis at $\alpha = .10, .05, .02, \text{and} .01$ (table 7a) for experimental group # 1 demonstrates the extreme significance shown by the vocabulary interventions for understanding connections between concepts (table 7a). In stark contradiction to the statistical analysis run for experimental group # 1, the analysis run for the other study groups (experimental # 2 and control) demonstrated no significant gains even at the $\alpha = .10$ level (tables 7b and 7c). Investigating the overall gains for the concept mapping assessment (as seen in table 6), experimental group # 1 had an average gain of 10.6%, whereas experimental group # 2 and the control group experienced score decreases by -2.61% and -3.44% respectively.
When students with an absence rate $\geq 50\%$ are entered into the data, it becomes evident that they create outlier points that affect the statistical significance of the study. With the addition of these students, none of the study groups experience statistically significant gains ($\leq \alpha = .10$) as seen from the Wilcoxon Signed Rank Test (table 8a-c). This is most likely attributed to the small sample sizes utilized in the study ($n = 15, n = 18, \text{and } n = 16$).

Experimental Group # 2 and the Control Group experienced overall average decreases in score amongst students participating in the study. A comparison of pre- and post test scores shows the majority of students either demonstrated no improvement or significant decreases in scores. Experimental Group # 1 is characterized by 87% ($n = 15$) of students maintaining their scores or showing improvement on the assessment whereas Experimental Group # 2 shows 72% ($n = 17$) maintaining scores or showing decreases in score. The Control Group is characterized by overall poor performance on the concept mapping assessment due to the fact that 81% ($n = 16$) of students demonstrated either no increase or significant decrease in score between pre- and post instruction.

The pre-instructional concept map scores varied slightly amongst the three study groups but this variation did not appear to affect the post-instructional assessment outcomes. Experimental Group # 1 showed an average pre-test score of 29.55 % and increased to 40.15% for the post-test assessment whereas scoring decreases are exhibited by both experimental group # 2 and the control (table 6) even though the pre-instruction assessment scores were similar. Although Experimental Group # 2 maintained the lowest pre-test score at 23.96%, it demonstrated a smaller decrease in score at -2.61% verses that
of the Control Group which began at 30.00% and experienced a drop of -3.44% to 26.66%.

The creation of the 15 question content sub-test arose from the fact that the 23 question GCI/OGT full test given to the students both pre- and post-instruction was not necessarily a true indicator of student understanding. The 15 questions were selected from the test that directly correlated to content being utilized during the plate tectonics unit. These questions were selected in hopes of finding a true indication of the students understanding of the content inherent in the unit. The questions were selected before any analysis of scores was initiated. The 15 questions included 12 GCI questions and 3 OGT (figure 4) questions. Students were expected to demonstrate scoring increases on this subset of questions due to the idea that these directly correlated to content being administered by the pre-set curriculum.

The removal of students with an absence rate ≥ 50% is an important factor when investigating the effectiveness of the vocabulary intervention on Experimental Group #1. When the 15 question content sub-test scores for the full population of students (n = 15) in Group #1 are examined it is found that 60% of students experience scoring increases and 13.3% scoring decreases from pre- to post-test (the remaining 26.6% maintained their original scores). When analyzed with the Wilcoxon Signed Rank Test it is found that a significant gain in scores can be assumed at the $\alpha = .10$ level (table 14a). The removal of the absent students from the population (n = 11) changes the rate of scoring increases to 63.6% and brings the rate of scoring decrease to 0% (the remaining 36.4% maintained their original scores). When analyzed with the Wilcoxon Signed Rank Test it is found that a significant gain in scores can be assumed at the $\alpha = .01$ level (table 13a).
demonstrating the effectiveness of the vocabulary intervention for the content being administered on plate tectonics.

Experimental Group # 2 experiences similar absence issues as group # 1 but does not demonstrate significant gains when the absent students are included or experience significant gains to the same level as that of Experimental Group # 1. With the full population (n = 18), it is found that the students scoring gains cannot be found to be significant even at the $\alpha = .10$ level. Experimental Group # 2 students (n = 18) show only 33% with scoring increases while 27% with scoring decreases (the remaining 38% maintained their original scores). Removal of the students with an absence rate $\geq 50\%$ allows for analysis that demonstrates significant scoring gains amongst the students (n = 16) to the $\alpha = .05$ level (figure 13a When absent students are removed from Experimental Group # 2 (n = 16), it seen that 43.75% of students increase in score whereas 12.5% decrease (the remaining 43.75% maintained their original scores). Although this was expected as students were exposed to the content, the level to which the gains are analyzed show that those experienced by students in Experimental Group # 1 were of greater significance than those in Experimental Group # 2.

The Control Group utilized in this study experiences similar results to the first two study groups but experiences higher scoring gains than experimental group # 2. With the full population (n = 16) it seen that significant gains exist in pre- vs. post test scoring to the $\alpha = .10$ level (figure 14c) where 62.5% of students showed scoring increases whereas 31.25% experienced scoring decreases. The removal of students experiencing an absence rate of $\geq 50\%$ allows for the control group (n = 15) to experience
significant gains at the $\alpha = .10$ level showing no overall change in the level of
significance with or without absent students.

The significance of the vocabulary intervention utilized in Experimental Group # 1
can be seen when comparisons of overall average group scores are analyzed. With
absences removed from all groups, Experimental Group #1 increases in score from 32.4%
to 42.67% (increase of +10.27%), Experimental Group # 2 increases in score from 36.6%
to 40.00% (an increase of +3.34%), and the Control Group increases in score from 34.1%
to 39.58% (an increase of +5.48%). The importance of these gains can be seen in
Experimental Group # 1 due to the fact that not only did this study group start with the
lowest average pre-test score, but ended with the highest post-test score. The +10.27%
average gain experienced by Experimental Group # 1 is nearly double that of the Control
Group and more than that for Experimental Group # 2. The effectiveness of the
vocabulary intervention for Experimental Group # 1 in relation to the content being
taught can be seen from the statistical data created from this 15 question content sub-test.
Although all study groups experience some scoring gains, it is evident that the most
impressive scoring gains were demonstrated by students in Experimental Group # 1 (table
12).

As with the previous assessment tools, rates of absence greatly affect the scoring
outcomes demonstrated by students in the Full 23 Question Test. The use of the Wilcoxon
statistical analysis (Figures 19a-c) demonstrated that none of the study groups
experienced significant scoring gains when the full populations were utilized. The
removal of the students experiencing an absence rate $\geq 50\%$ allowed for statistically
significant gains to occur for experimental group #1 and the control group to the $\alpha = .05$ level.

Experimental group #1 students demonstrated scoring gains amongst 60% of its students ($n = 15$) when the full population is analyzed and 72.7% of its students ($n = 11$) when the population is analyzed without the absent students. Experimental group #2 students experienced a scoring increase amongst 44% of its students ($n = 18$) when the full population is analyzed and an increase amongst 50% of its students ($n = 16$) when the absent students are removed. The students of the control group in this study experienced scoring gains amongst 56% of its students in the full population ($n = 16$) and 60% of its students ($n = 15$) when the absences are removed.

When average pre- and post-test scores are evaluated amongst the study groups for the full 23-question test, it can be seen that experimental group #1 not only experiences the greatest scoring increase but also obtains the highest post-test score (table 18). Experimental group #1 begins with an average pre-test score of 28.46% and ends with an average post-test score of 38.74% showing a +10.3% gain. Experimental group #2 experiences an average post-test score of 32.61% and an average post-test score of 35.60% (+3% gain). The control group students experienced an average pre-test score of 27.83% and an average post-test score of 34.78% showing a gain of +6.95%. The gains experienced by students in experimental group #1 are triple of those experienced by experimental group #2. The significant scoring increase experienced by students in experimental group #1 helps to demonstrate the effectiveness of the vocabulary intervention utilized throughout the study with experimental group #1 and its effect at
focusing on content comprehension. The scoring gains experienced by students in the
two other study groups can possibly be attributed to content exposure or teacher effect.

The 15 question GCI sub-test was utilized in this study so the results could be
compared to that of other studies that have used the GCI as an assessment tool. The
purpose of this study did not aim to address or compare other results therefore this data is
only being presented for future study use.

The data for the 15 question GCI sub-test followed similar patterns to that of the
full 23 question test where significant scores occurred in both experimental group # 1 and
the control group. With the absences removed from experimental group # 1 and the
control group, significant gains were identified to the $\alpha = .01$ level (table 26a-c and 27a-
c). Experimental group # 2 did not demonstrate any significant gains as measured to the
$\alpha = .10$ level.

The students of experimental group # 1 without absences (n =15) demonstrate that
66% of the students increase in score whereas 13% experience a decrease in scores.
When students with an absence rate $\geq 50\%$ are removed from the study (n = 11), it is
found that 72.7% of students have scoring increases while 0% of students experience a
scoring decrease (the remaining 27.3% maintained their original scores). Students of
experimental group # 2 did not exhibit any significant gains and it was found that of the
full population of students (n = 18), 33% experienced scoring decreases. When the
absent students were removed from experimental group #2 (n = 16) 25% of students
experienced scoring losses. The students of the control group demonstrated significant
gains on 15 question GCI sub-test. When the full population (n = 16) was analyzed, it
was found that 56.25% of students demonstrated scoring increases while 18.75% showed
scoring decreases. With the removal of absent students (n = 15), 60% showed scoring increases while only 13.3% had scoring decreases.

Investigating average 15 question GCI sub-test scores (table 24), it can be seen that experimental group # 1 posted an average scoring increase of +12.54%, experimental group # 2 an increase of +4.48%, and the control group a gain of +8.37%. It is evident from the statistical analysis utilized on this test that experimental group # 1 not only posted the largest scoring gain but also scored on average higher than students in experimental group #2 or the control group.

Conclusions

This study investigated the effectiveness of utilizing vocabulary interventions on ninth graders participating in a unit on plate tectonics. The use of quantitative assessment tools allowed for statistical interpretations to be made regarding the impact the interventions had on a student’s ability to not only learn the content but also establish deeply rooted connections about the concepts being delivered through varying instructional practices. The significant scoring gains achieved by students in this study receiving vocabulary interventions (experimental group # 1) can be viewed as a strong piece of evidence demonstrating the effectiveness of addressing vocabulary and language in the sciences. The students of experimental group # 1 not only showed the largest scoring gains but also demonstrated the highest post-test scores on every assessment.

The generalizability of this study to other populations is influenced by the unique character of the study population. The population from which this study sample was taken is characterized by relatively low test scores as seen from state-mandated
standardized tests (FIG 4). As a consequence, the effectiveness of the study should not be viewed in terms of final scores but rather in the gains achieved by the students from pre-to post-test. Experimental group #1 students participating in vocabulary interventions showed average scoring gains from 10.1% up to 15.1% on both multiple-choice and concept mapping assessments. These score increases demonstrate extensive gains in both content knowledge and conceptual understanding. Students from the other two study groups (experimental group #2 and the control group) showed scoring gains on the multiple-choice tests that may be accountable from exposure to the content throughout the unit. Alternatively, both groups demonstrated scoring decreases on the concept mapping assessment. Although these students were able to answer content-based questions about plate tectonics, they were unable to establish deeply rooted connections between the concepts being addressed as evident from their decline in scores from pre- to post-mapping.

The results from this study provide an interesting conundrum facing the educators of our children in regards to finding methods that truly gauge a student's understanding of the ideas and concepts being taught in the classroom. Public high school students in the state of Ohio must pass the Ohio Graduation Test (OGT) with a minimal score of 75% to qualify for graduation. Ohio views these passing students as scientifically literate, but a standardized test may not portray one's true conceptual understanding. The concept mapping assessment utilized in this study provided a mechanism for students to demonstrate their understanding of relationships between concepts. The maps created by the students allow them to show where similarities and connections exist amongst the concepts while demonstrating their knowledge of the overall content inherent in the unit.
The ability of a student to recognize that plate tectonics is not just a theory about how plates move but also to recognize that plate tectonics is an integral part of how the face of the earth is shaped shows that a student has gained conceptual knowledge of the process, not just a general understanding of the word. These maps allow teachers to investigate which students truly grasped the concepts and identify areas where misconceptions may exist. The ability of a student to demonstrate scoring gains on pre-map and post-map comparison is a strong indicator that increases in conceptual understanding were achieved. The vocabulary interventions experienced by students in experimental group #1 may have played a vital role in overcoming the language barrier existing in the sciences and allow these students to explore the concepts in greater detail.

The activities utilized in this study that helped provide the treatment to experimental group #1 may be beneficial if applied to other curriculum areas such as mathematics or humanities. A student needs to be provided with the tools to navigate their way through the heavy curriculum loads they are exposed to in school and the ability to work with the terminology may be a crucial piece of the puzzle for accomplishing this challenge. As with the sciences, each and every genre of study is laden with terms that must be understood before true conceptual understanding can take place. The use of vocabulary interventions in non-science courses may provide students with the opportunity to succeed in these classrooms as well.

The students exposed to vocabulary interventions (experimental group #1) experienced significant scoring gains that provide some insight into the effectiveness of the interventions. However, it must also be noted that the students exposed to lecture and book work (control group) experienced significant scoring gains on the multiple-choice
test as well. The teacher for the control group had over nine years of teaching experience whereas the teacher for experimental group #1 only had one year of experience. It would be interesting to have these two teachers switch places and then determine the effects of the vocabulary intervention in correlation with the teacher expertise and experience. If the vocabulary interventions are truly an effective tool at increasing conceptual understanding then the scoring gains should increase when the intervention is implemented by an expert teacher. In addition to switching teachers, the use of concepts maps in the vocabulary intervention may be better used as a learning tool rather than an assessment tool. If students were to create maps throughout a unit on plate tectonics, it may provide an even stronger opportunity for conceptual growth.

Several key findings deserve reiteration here, particularly with respect to significance of pre- versus post-test comparisons:

1. The vocabulary intervention utilized with experimental group #1 provided the students with the opportunity to increase their understanding of both content knowledge and conceptual understanding for the unit on plate tectonics.

2. The vocabulary interventions utilized with experimental group #1 demonstrated effectiveness for increasing student comprehension rates as seen through quantitative analysis of pre- and post test assessment scores.

3. Experimental group #1 posted the largest average scoring gains for the concept mapping assessment (+10.6%), the 15 question content sub-test (+10.27%), the
full 23 question GCI/OGT test (+10.3%), and on the 15 question GCI sub-test (+12.54%).

4. Experimental group # 1 achieved the highest post-test scores on the concept mapping assessment (40.15%), the 15 question content sub-test (42.67%), the full 23 question GCI/OGT test (38.74%), and on the 15 question GCI sub-test (39.10%).

5. The use of vocabulary interventions in a science classroom may provide students with the opportunity to increase content knowledge and conceptual understanding of any topic they are used in correlation with. The vocabulary interventions may provide for higher achievement on assessments.

6. The lecture based instruction delivered for the plate tectonics unit to the control group showed to be more beneficial than the instruction delivered to experimental group # 2. Scoring increases were higher for the control group and showed significant gains on the 15 question content test, full 23 question GCI/OGT test, and on the 15 question GCI sub-test.

7. The significant scoring increases demonstrated by the control group on the standardized tests may be accountable to teacher effect. The instructor for the control group had nine years of teaching experience and it is felt that the
knowledge and experience played an integral role in affecting the outcomes of the student’s assessment scores.

8. Although the students in the control group posted significant scoring gains on the standardized multiple choice test, they showed declines in conceptual understanding as demonstrated by their concept mapping scores (-3.44%). This shows that although they can recognize the content being utilized in the unit, they cannot establish deeper connections between content as shown by experimental group # 1 (+10.6%).

9. Students with absent rates of \( \geq 50\% \) played an integral role in determining the significance of the study. Students who missed 50% or tended to fall within the lowest performance level on the assessments.

References


Libarkin and Anderson, 2005. GSA presentation on Rasch Analysis.


Meagher, J. A. Multiple Meanings of Words -- Intermediate: A Comparison of Three Teaching Techniques. Paper presented to the International Reading

Nagy, W. E. Teaching vocabulary to improve reading comprehension / William E. Nagy. ERIC Clearinghouse on Reading and Communication Skills, 1988


Zeilik, M. “Concept Mapping”. Department of Physics and Astronomy, University of New Mexico.
Appendix A

DEMOGRAPHICS:
Please answer the following questions about your background.

**Gender**

**Birthdate:** Day_____ Month______ Year_______

**Racial Background:** ___White ___Hispanic ___Asian
___African-American ___Pacific Islander
___American Indian ___Other________

**Highest degree of:**

**Female Parent:**
___Elementary School
___some High School
___High School
___some College
___Bachelor’s Degree
___some Graduate School
___Master’s Degree
___Doctoral Degree

**Male Parent:**
___Elementary School
___some High School
___High School
___some College
___Bachelor’s Degree
___some Graduate School
___Master’s Degree
___Doctoral Degree

1.) Which of the following are actively contributing to the heat inside the Earth? **Choose all that apply.**

   (A) Gravity from the Sun
   (B) The concentration of the Universe's energy inside the Earth
   (C) Heat from the Sun
   (D) Radioactivity inside the Earth

2.) Rocks found in oceans can be _________. **Choose all that apply.**

   (A) Formed by animals
   (B) Made up of pieces of continental rocks
   (C) Formed by volcanic activity

3.) Which of the following can be caused by wind? **Choose all that apply.**

   (A) Movement of tectonic plates
   (B) Waves
   (C) Earthquakes
   (D) Mountain-building
   (E) Erosion
4.) Which of the following describes what scientists mean when they use the word “earthquake”.  
Choose all that apply.

(A) All earthquakes create visible cracks on the Earth's surface  
(B) When an earthquake occurs, the earth shakes at least once every 10 seconds for a period of at least 1 minute  
(C) All earthquakes damage man-made structures  
(D) When an earthquake occurs, energy is released from inside the Earth  
(E) When an earthquake occurs, the gravitational pull of the Earth increases

5.) Which of the following can greatly affect erosion rates? Choose all that apply.

(A) Rock type  
(B) Earthquakes  
(C) Time  
(D) Climate

6. Which is the best definition of a tectonic plate?

(A) All solid, rigid rock beneath the continents and above deeper, moving rock  
(B) All solid, rigid rock and loose dirt beneath the Earth's surface and above deeper, moving rock  
(C) All solid, rigid rock that lies beneath the layer of loose dirt at the Earth’s surface and above deeper, moving rock  
(D) All solid, rigid rock beneath the continents and oceans and above deeper, moving rock  
(E) The rigid material of the outer core

7.) Some scientists claim that they can determine when the Earth first formed as a planet. Which technique(s) do scientists use today to determine when the Earth first formed? Choose all that apply.

(A) Comparison of fossils found in rocks  
(B) Comparison of different layers of rock  
(C) Analysis of uranium and lead in rock  
(D) Analysis of carbon in rock  
(E) Scientists cannot calculate the age of the Earth
8.) The following maps show the position of the Earth’s continents and oceans. The o’s on each map mark the locations where volcanic eruptions occur on land. Which map do you think most closely represents the places where these volcanoes are typically observed?

Circle one:  

A. Mostly along the margins of the Pacific and Atlantic Oceans  
B. Mostly along the margins of the Pacific Ocean  
C. Mostly in warm climates  
D. Mostly on continents  
E. Mostly on islands
9.) On continents, where does most volcanic material come from?

A. Material comes from the Earth's center, which is completely molten.

B. Material comes from a molten layer near the Earth's center

C. Material travels from the Earth's center to a molten layer just beneath the surface, mixes with this molten layer and then travels to the volcano.

D. Material comes from the molten layer beneath the Earth's surface

E. Material comes from pockets of molten material beneath the Earth's surface
10. The figure below is a view of one-half of the Earth’s surface as seen from space today. The gray areas represent land, and the white represents water. Which of the other figures do you think most closely represents this half of the Earth’s surface when humans first appeared on Earth?

Circle one: A B C D

11. If a single continent did once exist, how long did it take for the single continent to break apart and form the arrangement of continents we see today?

(A) Hundreds of years
(B) Thousands of years
(C) Millions of years
(D) Billions of years
(E) It is impossible to tell how long the break up would have taken
12.) Scientists often talk about the Earth’s tectonic plates and their role in mountain formation, volcanism, and earthquake occurrence. Which of the following figures most closely represents the location of the Earth’s tectonic plates?

Circle one:   A    B    C    D

![Diagram of Earth's Tectonic Plates]

13.) Which of the following responses best summarizes the relationship between volcanoes, large earthquakes, and tectonic plates?

(A) Volcanoes are typically found on islands and earthquakes typically occur in continents. Both volcanoes and large earthquakes occur near tectonic plates.
(B) Volcanoes and large earthquakes both typically occur along the edges of tectonic plates.
(C) Volcanoes mostly occur in the center of tectonic plates and large earthquakes typically occur along the edges of tectonic plates.
(D) Volcanoes and large earthquakes both typically occur in warm climates near tectonic plates.
(E) Volcanoes, large earthquakes, and tectonic plates are not related, and each can occur in different places.

14.) Where are volcanic rocks found?

(A) Mostly on islands or in the ocean
(B) Mostly near the equator
(C) Mostly on the edges of continents
(D) On islands, in the ocean, near the equator, or on the edges of continents only
(E) Almost anywhere
15.) What does "density" refer to?

(A) How thick something is
(B) How quickly particles are moving inside something
(C) How much material exists in a space
(D) The amount of air contained within a material
(E) How slowly liquids move

16.) Why do tectonic plates move?

(A) The eruption of underwater volcanoes pushes the tectonic plates
(B) Currents in the ocean push against the tectonic plates
(C) Earthquakes push the tectonic plates
(D) Material is moving beneath the plates
(E) Magnetism moves the tectonic plates

17.) The map below shows the position of the Earth’s continents and oceans today. The gray areas represent land, and the white represents water. Which of the following best explains why the ocean basins look the way they do?

(A) Meteor impacts caused the ocean basins to form this way
(B) Ocean basins form as continents move
(C) The ocean basins formed in cracks that were created as the whole Earth cooled after its formation
(D) The ocean basins formed in cracks that were created as the whole Earth heated after its formation

18.)

Scientists believe that forces in Earth’s mantle move Earth’s crustal plates.
What do the arrows in the diagram represent?
A. ocean currents
B. gravity
C. convection currents
D. wind patterns
19.) The early development of the theory of plate tectonics was supported by which of these observations?
   A. matching fossils on the continents of Africa and South America
   B. glacier deposits far from existing continental glaciers
   C. thick sediment layers at the mouths of rivers
   D. sudden volcanic activity of long-dormant volcanoes

20.) ![Map of the Hawaiian Island Chain]

   The Hawaiian Islands are riding on the Pacific Plate as it moves northwestward. They are being formed as the plate moves over a hot spot in the mantle.

   Where is the next volcano likely to form?
   A. A
   B. B
   C. C
   D. D

21.) Some people believe there was once a single continent on Earth. Which of the following statements best describes what happened to this continent?

   (A) Meteors hit the Earth causing the continent to break into smaller pieces
   (B) The Earth lost heat over time and cracked, causing the continent to break into smaller pieces
   (C) Material beneath the continent moved, causing the continent to break into smaller pieces
   (D) The Earth gained heat over time and cracked, causing the continent to break into smaller pieces
   (E) Only a small number of people believe there was once a single continent, and it is more likely that the continents have always been in roughly the same place as they are today
22.) How far do you think continents move in a single year?

(A) A few inches  
(B) A few hundred feet  
(C) A few miles  
(D) Scientists do not have enough information to calculate the speed of continents  
(E) Continents do not move

23.) What did the Earth's surface look like when it first formed?

A. One large landmass surrounded by water  
B. All water and no land
Appendix B

Section 6: Worksheet 5 - Gondwana Puzzle

Puzzle Answer:

Instructions:
Fossils of Mesozoic plants and animals help scientists figure out how continents fit together during the Mesozoic Era. By matching the fossils zones shown in this puzzle, students can reassemble the continents to form a great southern supercontinent - called Gondwana. Give each student a copy of this sheet with the puzzle answer removed. Have students cut out the continents and color the patterned areas with colored pencils, crayons, or markers. Once colored, have students assemble the continents (Africa, India, Australia, Antarctica, South America, and Madagascar) together with tape and label each continent. The patterned areas (fossil plant and animal occurrences) along with the shape of each landmass will provide clues to how the continents fit together.

The Field Museum • Educator Guide: Part 1 • Section 6: The Permian
# Earthquake/Volcano Charting Data:

<table>
<thead>
<tr>
<th>EARTHQUAKES</th>
<th>VOLCANOES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude</td>
<td>Longitude</td>
</tr>
<tr>
<td>1.) 20.208°S</td>
<td>173.904°W</td>
</tr>
<tr>
<td>2.) 14.660°S</td>
<td>167.074°W</td>
</tr>
<tr>
<td>3.) 27.139°S</td>
<td>70.645°W</td>
</tr>
<tr>
<td>4.) 16.107°N</td>
<td>98.635°W</td>
</tr>
<tr>
<td>5.) 27.014°S</td>
<td>70.744°W</td>
</tr>
<tr>
<td>6.) 60.450°N</td>
<td>167.032°E</td>
</tr>
<tr>
<td>7.) 27.164°S</td>
<td>70.765°W</td>
</tr>
<tr>
<td>8.) 2.294°N</td>
<td>98.960°W</td>
</tr>
<tr>
<td>9.) 41.754°N</td>
<td>80.702°E</td>
</tr>
<tr>
<td>10.) 32.525°N</td>
<td>115.194°W</td>
</tr>
<tr>
<td>11.) 46.198°N</td>
<td>122.190°W</td>
</tr>
<tr>
<td>12.) 54.205°N</td>
<td>161.368°W</td>
</tr>
<tr>
<td>13.) 39.342°N</td>
<td>123.243°W</td>
</tr>
<tr>
<td>14.) 20.285°S</td>
<td>177.925°W</td>
</tr>
<tr>
<td>15.) 34.697°N</td>
<td>139.220°E</td>
</tr>
<tr>
<td>16.) 8.113°N</td>
<td>82.732°W</td>
</tr>
<tr>
<td>17.) 60.083°N</td>
<td>147.894°W</td>
</tr>
<tr>
<td>18.) 19.923°N</td>
<td>155.197°W</td>
</tr>
</tbody>
</table>
Using your notes from class today and the associated reading, create a sketch of the interior of the Earth that demonstrates:

1.) The Inner Core, Outer Core, Mantle, and Crust  
2.) How Convection occurs in the Mantle  
3.) Color the rising currents in Red  
4.) Color the sinking currents in Blue
Looking Inside Earth

Although the interior of Earth is only a few tens of miles beneath our feet, it is more difficult to reach than the surface of Pluto or even a nearby star! The deepest mines in the world are only three to four kilometers deep. The deepest well ever drilled only penetrates 12 kilometers into the interior. Since it is 6371 kilometers to the center of Earth, comparatively speaking, we have not even been able to get through the skin of the peach. Why haven't we gone any deeper? Even at a depth of 12 kilometers, the high pressures and temperatures of the surrounding rock cause the metal of even the strongest drills to quickly weaken and deform. At depths of 100 kilometers, rock flows like butter, and any hole we could form would quickly close.

Yet we have learned a good deal about Earth's interior using indirect means. Earthquakes and nuclear explosions generate sound (seismic) waves that travel through the entire volume of Earth. By measuring the intensity and timing of seismic waves at many locations on the surface of Earth, we are able to determine the density of materials at different levels in the interior. (The technique is a very sophisticated version of tapping on a barrel or wall to see if there is something inside.) By measuring the properties of different materials under high pressure in the laboratory, we can make a good estimate of the composition of different parts of Earth's interior.

From all these measurements, we find Earth's interior is layered something like the inside of a peach. The "pit" at the center of Earth is called the "core". It is made mostly of iron, some nickel, and about 10 to 15% of a less dense material, probably silicon, oxygen, or sulfur. The core itself is made of two concentric pieces, a solid inner core and a liquid outer core.

The large zone corresponding to the flesh of the peach is called the "mantle". The mantle is also divided into two solid parts, the lower mantle and the upper mantle. This division in the mantle may be due to small differences in the amounts of iron and magnesium or to differences in the nearness of atoms to each other. (As an analog, think of a piece of Styrofoam. If you crush it, the particles inside move closer together because you have closed the tiny open spaces inside, but it is still made of Styrofoam.)

Text Coding Key

*** Important Fact

Important / New Vocabulary

Confusing Information
Heat Flow

Common experience tells us that heat flows from hot to cold, so the heat in Earth's deep interior must be flowing somehow to the surface. Hot lavas and gases coming out of volcanoes are direct evidence of heat flowing out of Earth.

Both the measurements and simulations show that the hottest part of Earth's interior is the iron core. Part of the heat down there is actually left over from the fiery formation of Earth; part is from latent heat released by the freezing of liquid iron in the outer core onto the solid inner core, and part is (possibly) from the slow decay of naturally radioactive elements like uranium and potassium mixed in the core.

The core heats the bottom of the rocky mantle. The hottest rock near the bottom of the mantle becomes slightly less dense than the somewhat cooler rock above it, so buoyancy forces try to push the hottest rocks upward. Although the rock in the mantle is solid, the pressures and heat are so great that the rock can deform slowly, like hot wax. So the hot rock creeps upward through the cooler rock. As the hot rock rises, cooler rock flows downward to take its place next to the core, where it is heated and becomes buoyant enough to rise again later. The rising hot rock comes in contact with cold rocks near the surface of Earth where it gives off its heat, cools, and sinks again. Most of the rock in the mantle moves in this broad cyclic flow, indicated by the arrows in the figure. This zone, where rock is soft enough to flow, is called the asthenosphere.

(This means of heat transport--the cyclical movement of hot and cold material--is called convection. You can see an example of this in your kitchen by heating a pan of water to what is called a "rolling boil": hot water from the bottom of the pan rises up the sides, flows to the center, and sinks to the bottom again.)
Exit Slip for Monday

Due Tuesday

Name:

Directions: Use the word bank to correctly complete the vocabulary words below:

Word Bank

<table>
<thead>
<tr>
<th>Continental Drift</th>
<th>Convergent Boundary</th>
<th>Divergent Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earth-quake</td>
<td>Fault</td>
<td>Rift</td>
</tr>
<tr>
<td>Subduction</td>
<td>Tectonic Plate</td>
<td>Trench</td>
</tr>
<tr>
<td>Tsunami</td>
<td>Density</td>
<td>Lava</td>
</tr>
<tr>
<td>Hot Spot</td>
<td>Continental Plate</td>
<td>Oceanic Plate</td>
</tr>
<tr>
<td>Asthenosphere</td>
<td>Gravity</td>
<td>Igneous</td>
</tr>
<tr>
<td></td>
<td>Plate Tectonics</td>
<td>Ridge</td>
</tr>
</tbody>
</table>

S_b__uc__i__n    _a_lt    R__t
_a-F__or   S____ding    C__g__t_u__ary
__t__ic P__te    T__ch    _eous
P__e Tectonics    R__e    _a    _ma    H__ Spot
_ont__al Plate    _cano    Tsu__i    e__ity
__c Plate    _vection    Asthenos__
Gra__duction    _toni_ Pla__ench
Conti__al Dr__t    Di__gent Bo__ry
E__ake
Sentence Completion Exit Slip for Tuesday

Due Wednesday ______________________ Name: ______________________

Directions: Fill in each blank with the word from the word bank below that best fits the context. You may use each more than word once.

Word Bank:

Convection Asthenosphere Tectonic Plate(s) Density
Magma Mantle

Sentences:

1.) Convection currents of the melted material in the _________________ of the Earth cause slow movement of the ____________ ____________.

2.) Materials with less ____________ are pulled less by gravity.

3.) ____________ will have a stronger attraction for the cooled magma.

4.) The cooled ____________ condenses increasing the density.

5.) ____________ currents occur in the mantle of the Earth.

6.) A current of heat flowing from the core to the crust is a ____________ current.

7.) The convection current along the bottom of the crust causes the moving of the _________________ ____________.
Exit Slip for Wednesday
Due Thursday

Directions: Fill in each blank with the word from the word bank below that best fits the context. You may use each more than word once.

Word Bank:

Continental Drift    Ridge    Density    Magma    Sea-Floor Spreading

Asthenosphere      Gravity    Convection    Trench    Tectonic Plate

Sentences:

1.) Sea-Floor Spreading is the idea that sea floor spreads and __________ moves into the cracks.

2.) ______________________ __________________ is the theory that the continents of the Earth fit together and drifted apart.

3.) Materials with less __________ are pulled less by ________.

4.) Fossil, climate, glacial, and geological evidence provide evidence in favor of ______________ ____________.

5.) Scientists discovered the youngest rock in the sea floor to be at or near the ______________.

Determine the vocabulary word being defined:

1.) __ _ __ e_ t_ o_ : A transfer of heat in gas or liquids by currents.

2.) __ a _ i _ _ : The attraction of the Earth for any object.

3.) __ d _ _ : An area in the ocean basin where new ocean floor is formed.

4.) _e_ t_ _ _ c _ _ _ t _ : A moving plate of the Earth’s crust and upper mantle.

5.) __ t _ e _ o _ _ _ _ r _ : Plastic-like layer of the Earth’s mantle upon which plates move an shift.
**Exit Slip for Thursday**

_ _______: A _______ in the _________ crust along which rocks move when broken

_ _____ tion: When ____________and __________ plates collides, and the oceanic plate moves below the continental plate

Te___c P___te: A moving _______ of the Earth’s ______ and upper mantle

D___gent Boundary: A _______ where two _______ are moving apart.

_ ____ch: A steep sided ________ formed on the sea floor when one ______ moves beneath another plate

A___osphere: Plastic-like layer of the Earth’s mantle upon which plates move

__ eta___: Rocks formed when existing _____ undergo high temperature, high _________, or other chemical changes

_ ___ity: The attraction of the ___ for any object

Sea ___ Spreading: A process of plates moving _______ and allowing ______ to move upward and form new sea floor

_ ___ity: The _____ per unit ______ of a given substance

Lava: The ______ that flows onto the surface of the earth

___ Plate: A rigid, thin layer of the Earth’s Lithosphere considered to have a higher density than
continental plates

**Hot _______:** Areas near the center of _______ where volcano activity occurs

**Continental _______:** a rigid thin layer of the Earth’s _______.

**Igneous:** Rock that forms when magma cools and hardens

**_ _ _ _ _no:** An opening in the Earth’s _______ through which lava flows or erupts

**Rift:** A long narrow _______ formed as two plates move apart

**_ _ _ _ nami:** A sea wave formed over the focus of an _______

**M_ _ _a:** The melted rock under the Earth’s surface

**_ _ _ vection:** A transfer of _______ in gas or liquids by currents

**Ea_ _ _ _ _ ake:** Vibrations formed when rocks break along a________.

**R_ _ _ e:** An area in the ocean basin where new ocean floor is formed

**C_ _ _ inental Drift:** A hypothesis that states ____________ were once a single ____________ that broke apart and drifted away from each other.

**Convergent Boundary:** A _________that forms when two plates come together.