2006

Measuring Student Understanding of Density, with Geological Applications

Emily L. Klingler

MEASURING STUDENT UNDERSTANDING OF DENSITY, WITH GEOLOGICAL APPLICATIONS

By

Emily L. Klingler

B.A. Boston University, 2003

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Teaching

The Graduate School

The University of Maine

May, 2006

Advisory Committee:

Stephen A. Norton, Professor of Earth Sciences, Advisor

Peter O. Koons, Professor of Earth Sciences

John R. Thompson, Professor of Physics and Astronomy

LIBRARY RIGHTS STATEMENT

MEASURING STUDENT UNDERSTANDING OF

DENSITY, WITH GEOLOGICAL APPLICATIONS

By Emily L. Klingler

Thesis Advisor: Dr. Stephen A. Norton

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Teaching
May, 2006

Density is an important physical concept in all of science. Substantial education research on the teaching and learning of density has occurred in physics, chemistry, and biochemistry; there has been little education research about density in the earth sciences. This study investigates the extent to which targeted instruction influences student understanding of density, and the frequency and accuracy with which students apply it in explanations of earth science phenomena.

All data for the study were collected from students in The Environmental Geology of Maine (ERS 102) at the University of Maine during the spring and fall 2004 and spring 2005 semesters. Responses to three earth science-related post-course questions, from the final examination in spring 2004, served as baseline data for the study (n = 66), as well as a trial run for assessment questions. Based on the fall 2004 data, a pre-/post-course assessment on density was developed for targeted instruction. The assessment contained three questions about density in an earth science setting, and seven more general questions on mass, volume, and density.

The targeted instruction was implemented for two experimental groups (n_{total} = 97) in fall 2004 and spring 2005. The pre-course assessment determined students' prior

knowledge and understanding of the concept of density upon entering the class. An inquiry-based laboratory exercise, focused on density, was designed for use as the intervention in both experimental groups. The exercise centered on students determining the density of homogeneous and heterogeneous substances. The post-course assessment, identical to the pre-course assessment, measured students' understanding and performance after the new density exercise.

Chi-squared analyses of pre-/post-course assessment results for both experimental groups revealed three findings: (1) responses to approximately one third of assessment questions improved; (2) responses to approximately one third of assessment questions remained consistent or showed some decline in level of understanding; and (3) responses to approximately one third of assessment questions had insignificant differences. Overall, the density laboratory exercise appears to have positively influenced student understanding of density in and out of an earth science context, as measured by the pre-/post-course assessment. However, there is much room for revision to both research instruments.

DEDICATION

This thesis is dedicated to my parents, Eileen and Garrett Klingler, who have always encouraged me to follow my heart.

ACKNOWLEDGEMENTS

Thank you to everyone in the M.S.T. program and the Department of Earth Sciences for all your help and support! Special thanks to Jeff Owen for all his advice, encouragement, and energy! I probably would not have enrolled in the program if I had not first run into him. Many thanks go to the chair of my committee, Dr. Stephen Norton. Without all of his help and advice, this project never would have gotten off the ground! His infusion of energy, right from the start of the project, was ongoing and kept me working as hard as I could. Many thanks as well to the other members of my committee: Dr. John Thompson, Department of Physics and Astronomy; and Dr. Peter O. Koons, Department of Earth Sciences. Their different perspectives on my project helped shape its outcome from many angles, making the final work infinitely better!

Dr. Phillip A. Pratt deserves special thanks for all the advice he offered on the statistical tests necessary for our data. Though he was not a member of my committee, he freely gave his time and energy, and this project would not have been as complete without his expert advice.

Finally, thank you to the Graduate School for helping me through the last stages of my research and writing. The Summer Research Grant allowed me to focus my energy on this project and finish it on schedule.

TABLE OF CONTENTS

LIST OF TABLES

Background

Complex processes characterize earth science topics. Each process, however complicated, is built upon basic principles. For example, mantle convection at its most fundamental level is a result of density variations within the mantle. Density difference is the fundamental driving mechanism for many other earth science processes, including subduction and accretion of terrains, convection in active zones of permafrost regions, and thermo-haline oceanic circulation.

A student can learn about an earth science process in terms of what it represents. But without a solid understanding of why the mechanism exists, the student may leave the course with a superficial or incomplete understanding of that phenomenon. Without explicitly addressing the concept of density and its relation to many earth science processes, a student may not gain a deep understanding of the processes.

In a physical science study involving 8^{th} graders, Smith et al. (1997) noted that students must "be aware of their own ideas" that they bring into a class, in order to improve upon or change those ideas over the course of the class. Hannula (2003) noted that asking earth science students to formulate hypotheses prior to conducting measurements is a great way to make students' ideas explicit in the classroom. Hannula asked students for feedback on the laboratory exercises, which had been modified to include students' hypotheses. Students' self-assessment surveys showed that most students found the revised exercises produced an improved understanding. Furthermore, Smith et al. (1997) showed with 8^{th} grade physical science students that a laboratory curriculum, when revised to incorporate students' ideas and explore their questions, produces more significant increases in understanding than traditional menu-driven laboratory curricula. Their study found 85% improvement in student understanding of density, mass, and volume, compared to only 35% improvement with the standard physical science curriculum!

Education research on the teaching and learning of density has occurred primarily in physics, chemistry, and biochemistry (Yeend et al., 2002). All studies found that students at any grade level consistently have trouble understanding mass, volume, and density. Driver et al. (2003) reported for 11-year-old students that “over 80 per cent had misconceptions about volume which could present serious difficulties for understanding density.” Driver et al. (2003) found that children confuse density with weight.

From interviews with primary, intermediate, junior high, and college students, Stepans et al. (1986) “found that a large number of college students said that when you crumple an aluminum sheet, you have made it heavier,” showing confusion between mass and volume. Driver et al. (1985) noted that the concept of mass is confusing for students, possibly because weight is commonly used in place of mass. Some students have the idea that mass varies for an object, though weight is the measure that would actually vary in different environments. Clearly, if students harbor confusion about mass or volume, they will have trouble understanding the idea of density.

There have been few education research studies specifically on density in the earth sciences. McKinnon (1971) noted that college-level earth science textbooks commonly discuss topics with the assumption that the reader has a thorough understanding of density, though many college students have an insufficient understanding of the concept. In a study of college students' understanding of sinking and floating experiments, he found that one of four students had trouble understanding volume.

Many studies, particularly in physics, have attempted to measure student understanding of density in the context of buoyancy (McKinnon, 1971; Hewson et al., 1983; Loverude et al., 2003; Heron et al., 2003). Using student interviews, Libarkin et al. (2003) found that many students believed water was necessary to sort objects according to density. They noted that many students harbor alternate conceptions about density, because "standard density experiments involve objects floating and sinking in water…and many students mistake buoyancy-related phenomena for characteristics of density."

Common student difficulties in understanding and using density include:

- alternate conceptions about volume;
- confusion between mass and volume;
- confusion between mass and weight;
- confusion between weight and density; and
- alternate conceptions about sorting material by density.

The concept of density underlies many topics covered in the University of Maine's ERS 102, The Environmental Geology of Maine, which many college students complete as a fulfillment of a general education requirement. On average, the course consists of 90% general education students, 5% earth science majors/minors, and 5% other science students.

In the spring 2004 semester, density was highlighted in many lectures as an important concept underlying earth science processes, and the students had the opportunity to conduct a short density exercise in laboratory section. This laboratory exercise was held well after the start of the course and many density-influenced topics had been covered in lecture, though density was rarely explicitly connected to any of those topics, either in lecture or laboratory. Upon completion of the laboratory exercise, many students remained uncertain about the properties of mass, volume, and their ratio (density) as properties of a substance. Responses on the mid-term and final examinations for the course showed that many students could not demonstrate the connection between density and the earth science topics covered in the course.

We recognized that density in relation to earth science processes was an area of difficulty for many students. Consequently, the concept of density was selected for a two-year education research study. The study includes a baseline group (n = 66) and two experimental groups (n = 97). The baseline group was instructed using a traditional lecture and laboratory curriculum, and included a post-course assessment; the experimental groups were instructed using a modified curriculum, targeting density as an important concept. The modified curriculum included pre- and post-course assessments, and a specially designed, inquiry-based laboratory exercise on density, in order to investigate the extent to which targeted instruction influences student understanding of density, and the frequency and accuracy with which students apply it in explanation of earth science phenomena.

Methods and Instruments

Baseline data collection

Three earth science-related questions (Appendix C) were developed for the study, and were test-run by earth science graduate students. Their third-party perspectives on the appropriateness and clarity of the questions, as well as the type of responses we could expect from the questions, aided our refinement of the questions before implementation for the study. The three questions then appeared on the final examination in ERS 102, spring 2004, and served as baseline data for the study (n = 66). The baseline data only consists of post-course assessment responses for the entire class (i.e., population data). No pre-course assessment was administered to the baseline group. Thus, individual students' performances in the group were not tracked over time.

A short density laboratory exercise had been included in the curriculum for students in the baseline group (hereafter "Group 1"), but it was procedure-driven and very short. The exercise was designed to be a refresher on mass, volume, and density, rather than an education research tool.

The timeline and details of each data group for the study are in Table 2.1. The three questions administered to Group 1 were included on the pre-/post-course assessment on density, which were developed for the experimental groups (Group 2 = fall 2004; Group 3 = spring 2005). The assessment given to Groups 2 and 3 also included seven questions on density, mass, and volume (Appendices D, E).

Table 2.1: Timeline and details of data groups for the study

ERS 102 Semester	Title	Pre-course assessment	Post-course assessment	Function
Spring 2004 n = 66	Group 1	No	Yes- questions 1 to 3	Development of baseline data and contextual density questions
Fall 2004 n = 32	Group 2	Yes	Yes	Assessment of change in understanding from new laboratory exercise.
Spring 2005 n = 65	Group 3	Yes	Yes	Assessment of change in understanding from new laboratory exercise.

Experimental data collection

The targeted instruction was implemented for Groups 2 and 3 in all laboratory sections of ERS 102 in fall 2004 and spring 2005, which were taught by two different faculty members. There was no control group for the two experimental data sets for two reasons. (1) There were only two sections of ERS 102 in 2004, and four sections in 2005. Thus, using some of those sections for a control group would have yielded too small experimental data sets. (2) The professors believed that all students should receive the same lessons, so that no student had an advantage over another in the course.

For both experimental groups, the pre-course assessment determined students' prior knowledge and understanding of the concept of density upon entering the class. The inquiry-based laboratory exercise, focused on density, was designed for use as the intervention in both experimental groups, and was completed early in the semester by students. The exercise (described below) centered on students determining the density of homogeneous and heterogeneous substances. The post-course assessment, identical to the pre-course assessment, measured any change in students' understanding and performance as a result of the new density exercise or any other influence in the course. Finally, every student's performance on the pre- and post-course assessment was tracked with participants' student identification numbers. In this way, we knew that not only x% of the group improved their level of understanding of density from pre- to post-course, but y% of the students who originally responded at some level of understanding improved their level of understanding on the post-course assessment.

An informed consent form (Appendix B) was given to all students prior to administering the pre-course assessment, which explained that participation in this study was voluntary, would not affect his/her grade, and that s/he could opt out of the study at any time. The laboratory instructor also reviewed this information with the students, prior to administering the pre-course assessment, and stressed that all students complete the assessment to the best of their ability and leave no questions unanswered. The students were informed that the assessments were not given course credit, i.e., there was no extra credit allotted for participation. The pre-course assessment began with a permission question, asking if the student were willing to participate in this study. If a student decided not to participate, their assessment was not added to the data set for the study, but was retained for the semester so that the professor could look at the student's responses. The post-course assessment differed, in that it did not have a permission question.

Group 2 received the first iteration of the pre-course assessment (Appendix D) at the first laboratory meeting. Post-course assessments were administered during the last laboratory meeting of the semester. The pre- and post-course assessments were nearly identical, with some minor revisions to the post-course version, as discussed below.

The assessment consisted of two parts. Part I included three questions relating to density in an earth science setting. Density was not explicitly mentioned in Part I, because it was designed to measure how frequently, and how well, each student could identify and discuss density differences as the underlying cause for the process in each question. The idea was that students would invoke the idea of density on their own, with no prompting from the wording of the question. The students did not know that the earth science contextual questions were driving at the concept of density, but would discuss density if they felt it necessary. The title of the assessment did not even hint that density was the subject of the test, and instructors introduced the assessment simply as important to the class and to our study. The three scenario questions asked for a full open-ended explanation of the students' reasoning. Questions 2 and 3 of Part I were revised from the pre-course to the post-course version (Appendix E). Wording in question number 2, choices "a" and "b," were revised to be more clear and accurate: a plate is not "forced downward" but rather "sinks." Question number 3 was also revised: the well is not drilled into the cliff, but rather "parallel to the cliff."

Part II of the assessment dealt with density more abstractly. This portion consisted of seven questions, all explicitly asking about the mass, volume, and the ratio of mass to volume, or density, of an object. Each question attempted to approach density, mass, or volume in a slightly different manner, with the idea that all questions would help to pinpoint students' level of understanding, i.e., triangulation. For example, if only one question was asked on the density of an object and a student gave the correct answer, we might believe that the student had a complete understanding of the concept. But, if we asked about density four different times, and in four slightly different ways, we might tease out students' level of understanding, given their levels of response on all four of those questions.

Questions 4 to 7 were adapted from a survey used by Yeend et al. (2002) and Loverude (2004). Questions 4 to 6 dealt with three portions of a straight, uniform board, labeled A, B, and C. These initial questions were randomly ordered to avoid leading the students to think back on the density equation, i.e., that density = mass/volume. For this reason, we did not ask about mass, and then volume, and finally density. Volume was addressed first, then density, and finally mass. In order to reduce confusion over the portrayed size of the three board segments in these questions, text was added on the post-course assessment: "A is shortest, C is longest."

The second iteration of pre-/post-course assessments (Appendix E) was administered in spring 2005 (Group 3), also at the first and last laboratory meetings, respectively. The students received an informed consent form, and were asked to be a part of the study with the permission question at the top of the first page of the pre-course assessment.

Design of the Assessment

The assessment was specifically designed with distracters based on the misconceptions research discussed earlier:

- Alternate conceptions about volume: In addition to one question, 4, devoted solely to the volume of a rectangular solid, other questions incorporated the concept of volume. Questions 4, 5, and 6 were interrelated: if a student could not answer the volume question correctly, or if the student did not understand how volume relates to density, he may get questions 5 and 6 incorrect. Additionally, an understanding of volume was necessary in reasoning through questions 8, 9, and 10, which all dealt with rectangular solids of different sizes on balance beams. If a student did not understand the relation of volume to mass in terms of determining density, he would get these questions incorrect. Furthermore, if a student believed that larger objects were denser, he would respond incorrectly on all density questions in Part II of the assessment: 5, 7, 8, 9, and 10.

- Confusion between mass and volume: Not addressed with this assessment.
- Confusion between mass and weight: Not addressed with this assessment.
- Confusion between weight (or mass) and density: A student with this misconception would find choices he would agree with on all density questions in Part II of the assessment: 5, 7, 8, 9, and 10. For example, on question 7, which deals with the density of a diamond chip compared to the density of the original diamond, the student would say that the chip is lower in density than the original diamond, because the chip is smaller than the original.
- Alternate conceptions about sorting material by density: Not addressed with this assessment.

<u>Intervention: The Laboratory Exercise</u>

Every laboratory section participated in the inquiry-based laboratory exercise, which targeted the concept of density. I was present and involved in the instruction for all laboratory sections. The exercise consisted of two parts. Part I of the laboratory exercise was not earth science context specific, whereas Part II was (Appendix F).

Part I of the exercise asked students to work with Silly Putty™, bbs, or a Silly Putty™-bb mixture. There were two groups for each material, with students randomly assigned to the groups to avoid cliques or academic advantages among groups. Each group was asked to decide upon a procedure to determine the density of an irregularly shaped object, and then use their procedure to determine the density of their assigned material. Each group then made a five-minute presentation of their methods and results, with density measurements from all groups tabulated on the board. Each student was then asked to "summarize the findings for the determination of the density of the Silly Putty™, the bbs, and the Silly Putty™-bb mixture in the form of three 'rules,' and include these rules and an explanation of each one in your laboratory write-up." These rules were used to qualitatively describe the density of homogenous and heterogeneous materials, and were applied to Part II of the laboratory. The rules students deduced for the materials were as follows:

1. The density of a small piece of Silly Putty™ was the same as the density of the entire piece of Silly Putty™. This is because the Silly Putty™ is a homogeneous substance, i.e., the density is the same throughout the material.
2. The density of a small number of bbs was the same as the density of a large number of bbs. This is because the bbs were made of a homogenous substance, i.e., the density is the same throughout the material.

3. The density of a small piece of the Silly Putty™-bb mixture was not necessarily the same as the density of the entire piece. This is because the Silly Putty™-bb mixture is a heterogeneous substance, and the density may vary throughout the material.

In Part II, each group was given three samples each of the same rock and constituent minerals, and was asked to predict which rule from Part I was applicable to each rock and mineral. Groups then determined the density of each sample. This method helped students to understand that the rocks they were working with were heterogeneous composites of minerals, and that minerals are generally homogenous. This section also served as an introduction to the next laboratory exercise of the semester: igneous rocks and rock-forming minerals.

Our laboratory exercise was designed to address the research question: To what extent does targeted instruction influence student understanding of density, and the frequency and accuracy with which students apply it in explanation of earth science phenomena? The goals of the laboratory exercise were:

1. To improve students' understanding of density: what it is, how it is calculated, and what density represents for an object.
2. To give students an opportunity to determine the density of a common substance (water) and talk about variations between a measured and a theoretical density.
3. To teach students how to find the density of an irregularly shaped object (i.e., they cannot simply look up a volume equation for the shape).

4. To allow students a chance to test variation of shape against mass and volume.
5. To introduce proper laboratory techniques, experimental variability, the importance of replication, and sources of error.
6. To help students become acquainted with dynamics of group work, and to give students the opportunity to practice presenting to a group.

This targeted instruction was designed to meet each goal in the following ways (numbers below correspond to goal numbers above):

1. The first question in the laboratory exercise asked students to reflect on and discuss density with their group members. In the first iteration of the density exercise, each student was asked explicitly to write a sentence describing what density is without using "over," "per," or "divided by." It was hoped that students would think about density in a more qualitative way, instead of simply listing the density equation, which they may have memorized. Being able to recall an equation does not necessarily imply understanding! Students struggled with this question, and many spent too much time on it. Furthermore, the stipulations in the first question did not seem to aid students in thinking about density, because many students were confused about density at the beginning of the exercise anyway. In the revised version of the exercise for spring 2005 (described below), the first question changed, asking students to articulate their understanding of density. This revised approach seems to have been a better way to ease students into the exercise.

Smith et al. (1997) and Hannula (2003) noted that students' ideas must be made explicit in the classroom, in order to improve upon or change those ideas over the course of the class. Both versions of the density laboratory exercise asked students to formulate predictions, in order to get students' ideas down on paper. This way, students were committing themselves to an idea, which they could then test for validity. Many students were wary of predictions, fearing that they would be wrong. The instructors set students' minds at ease, and announced that their hypotheses would not be graded for correctness, but on whether they were included in the laboratory write-up. After performing their laboratory procedures, students were asked to evaluate their predictions, discuss whether they were correct, and suggest possible reasons for any disagreement between the prediction and results.

2. All densities determined by the groups were tabulated on the board and discussed by the class, which was a visual way to illustrate variation of experimental procedures. For example, each group had been asked to find the density of water, and each group's findings differed from those of other groups, as well as from the theoretical density of water, which most students seemed to remember should be $1 g/cm^3$. Each group discussed their measurement techniques in their presentation. Techniques were discussed by the class; instructors were careful not to single out any one group.

3. Students were given the materials to do this exercise. The list of equipment hints at how to find an object's volume by water displacement. Commonly, at least one person in each group was familiar with this method of volume measurement and could show the others in the group the technique. Monitoring and suggestions from the instructor helped students unfamiliar with finding volume in this way.

4. Students were allowed to experiment with the Silly Putty™ to determine whether various shapes and sizes affected density. Students given the bbs or the Silly Putty™-bb

mixture could benefit from the Silly Putty™ groups' presentations on shape and size variations, because the groups with the Silly Putty™-bb mixture were asked not to manipulate their sample more than dividing it into pieces (it had a very heterogeneous density from one end of the piece to the other, depending on the number of bbs in each end).

5. See explanation number 3.

6. Instructions explicitly asked each student to discuss results with the other members of their group. The instructor observed each group in turn, making sure the group members were communicating with each other and not working alone.

Each group presented their procedure and findings for approximately five minutes. This allowed peer review of each group and provided an opportunity for revision, if necessary. For example, some groups had used the incorrect formula for density or had a slightly incorrect procedure. The peer review and comparison to other groups' techniques were very worthwhile.

We decided that our density exercise used in the fall of 2004 (Group 2) was still too procedure-driven. Students were explicitly told a procedure to use for determining the density of their material. This component was in the exercise to assure sufficient time for each group to explore variation in size, shape, and/or number of material, and the corresponding densities. Observation of student groups working on the exercise in the fall of 2004 suggested that students could make these observations on their own. Perhaps if students' ideas were incorporated into the exercise, they would take more interest in the exercise and learn more from it. The structure of the revised inquiry-based laboratory exercise was in the style of *Physics by Inquiry* (McDermott et al., 1996), curriculum materials designed to teach K-12 pre-service teachers physical science as a process of (guided) inquiry.

Smith et al. (1997) showed that laboratory exercises, revised to incorporate students' ideas and explore their questions, produced far greater improvement in student understanding of density, mass, and volume than traditional curricula. With that study in mind, our density exercise was revised to be more of a guided inquiry exercise. Groups were given materials and a goal: to determine the density of the material they were given. Groups were free to design their data tables according to what they felt they should measure or test, and instructors checked tables to be sure that students included sections for recording the basics: mass, volume, and the calculated density of the material.

In the revised exercise, students were guided by questions from the instructor. For instance, if one group found the density of their piece of Silly Putty™ and claimed that they were done, the instructor might pull off a piece of the Silly Putty™ and ask the group what the density of that smaller piece was, and how they could be sure. This helped lead students to formulating their "rules" about their substance.

Part II of the laboratory exercise was revised slightly from Group 2 (Fall 2004) to Group 3 (Spring 2005). An extra-credit question on the density of pumice was added as one final experiment and thought question. Students were very surprised to find a rock that floated on water, and were asked, "What can explain the differences in behavior between the pumice and the rocks/minerals." Most students could see that the pumice had many air pockets and therefore likely had a lower density than the compact rocks and minerals they had tested.

Results

The research question was: To what extent does targeted instruction influence (1) student understanding of density and (2) the frequency and accuracy with which students apply this understanding to explain earth science phenomena?

Results from Group 1 (baseline group) are compared to results from Groups 2 and 3 (experimental groups) on the contextual questions (Part I, questions 1 to 3) of the post-course assessment, in order to address the research question. The post-course assessment results for Groups 2 and 3 should show higher levels of understanding than Group 1, if the targeted instruction positively influenced student understanding of density in an earth science setting.

The research question, addressed using performances on pre-/post-course assessments for Groups 2 and 3, breaks down as follows:

1. Was there significant increase in understanding of density from pre-course to post-course assessments?
2. Were students' responses consistent on questions 5 and 7?
3. Were students' responses consistent on questions 8 and 9?
4. Were students' responses consistent on questions 4, 5, and 6?

Numbers 2 though 4 refer to the triangulation of density questions: repetition was intentionally built into the assessment, to better discern a student's level of understanding. Performances on non-contextual density assessment questions (Part II of the assessment) are compared with one another, with exception to question 10, which we determined to be a confusing and poorly written question (described below).

Data analysis

The assessment contained extended response and multiple-choice questions, except for questions 8 and 9, which were multiple-choice. Students were graded on the level of understanding they demonstrated in their extended responses. In the case of an unclear or insufficient response, students were not given the benefit of the doubt, but were assumed to know only what they had shown on the assessment.

A rubric for assigning the level of understanding was developed for each question following Wiggins and McTighe (1998). Below is an example of one question from the assessment, and its rubric (Appendix H has all rubrics). Examples of each level of understanding are also shown.

All rubrics had six levels of response (listed in Table 3.1). Blank responses to questions were automatically graded at level 0. For assessment questions with multiple-choice portions, a student's response could fall at levels 0 through 3 without regard to whether his multiple choice response was correct; only correct multiple-choice responses were assigned at a rubric level of 4 or 5, depending on the quality of the corresponding explanation. A student whose answer to question 1 showed a sophisticated level of understanding received a rubric score of five for that question. The same student might have demonstrated an in-depth level of understanding on question 2, receiving a rubric score of four for that question, and so on for the eight extended response items.

"Question 1: What process in Earth's mantle is thought to cause plate tectonics? Explain the process fully."

Table 3.1: Rubric for Question 1

Explanation	Example
5. Sophisticated: convection may or may not be named, with a thorough explanation of convection (temperature and density differences cited).	"Convecting and reheating of earth's surface- material gets pushed down because of density and reheats." (the student has drawn convection currents in the mantle and a subduction zone to supplement her explanation.)
4. In-depth: convection may or may not be named, with a sufficient explanation of convection (involving temperature differences); going beyond what is obvious or what was explicitly taught.	"Convection currents: currents of hotter magma circulate up, cooler magma down. Currents at top of magma create forces which move tectonic plates."
3. Developed/superficial: convection/currents named and/or cells drawn with related but superfluous information and/or a relevant portion of the rock cycle explained; mechanisms for convection insufficiently explained.	"Forget the name but the mantle flows in a circular pattern causing the plates to almost float above and that floating leads to movement."
2. Intuitive: an incomplete/incorrect account, with apt and insightful ideas; account has sweeping generalizations or limited support; may cite an incorrect mechanism.	"Earthquakes and volcanoes in the core erupting up to the mantle causes separation and formation of giant mass."
1. Naïve: a superficial account; more of a restatement of the question than an explanation; a fragmentary sketch or glib generalization; more of a borrowed idea than a theory; cites an incorrect mechanism.	"The movement of the plate that constitute the Earth system."
0. No response	

Each rubric level was somewhat subjective and therefore difficult to quantify. The levels that were the most difficult to distinguish between were levels 4 and 5. All other rubric levels were more clearly identifiable in students' responses. However, at times, a response fell between rubric levels. In such cases, the responses were compared to other responses at the levels in question, and, at times, the rubric was revised to accommodate unforeseen responses. These changes certainly introduced error into the study.

The use of rubrics, in which responses are classified according to rank, yields ordinal data. Such data does not yield equal gradation between levels. Thus, parametric statistical tests such as a t-test, or any other test designed to manipulate discrete numbers (interval/ratio data), could not be used (Coladarci, 2004).

An appropriate test for ordinal, bivariate frequency data is the Chi-Square test of independence (Coladarci, 2004). This test uses a cross-tabulation to determine whether the responses on the post-course assessments were independent of responses on the pre-course assessments. A cross-tabulation allows the freedom to analyze one set of scores against another set of scores, for the same question pre- to post-course, or for scores from two different questions. For example, in order to determine how well each student understands density, we used a series of cross-tabulations to compare responses to each density question to the others. The level of significance (alpha) for all data analyses is 0.05. All statistical analyses were performed using the SPSS Student Version 11.0 for Windows™.

Population data: Comparing post-course data between groups

A summary of performances on assessment questions 1 to 3 (Appendix C) among all groups is in Table 3.2. Group performances were compared because individual students were not tracked in Group 1. Data for Groups 2 and 3 in the remainder of this study are for individuals whose performances on assessments were tracked from pre- to post-course, using student identification numbers. Cross-tabulations of performances are discussed only for the statistically significant data sets (tables corresponding to non-statistically significant results are in Appendix I).

Table 3.2: Significance of results on post-course assessment questions 1 to 3 among all groups (significance at $\alpha = 0.05$)

Groups	1 and 2			1 and 3		
Question	1	2	3	1	2	3
	no	no	yes	yes	yes	no

Performances on post-course assessment question 3 were significantly different between groups 1 and 2 (Table 3.2). The largest differences between group percentages occur at rubric levels 1 and 2 (Table 3.3a). Forty-nine percent of Group 1 scored at level 1, compared to only 19% of Group 2 scoring at that level. Group 1 was outperformed by Group 2 at level 2 (14% compared to 31%, respectively). Additionally, Group 2 had more students scoring at levels 3 and 4. Both groups had similar percentages of students scoring at level 5. These data indicate that Group 2, the experimental group, demonstrated better overall understanding on question 3 than Group 1, the baseline group.

Table 3.3b and analogous tables that follow give the numerical value of the Chi-Square analysis for previous tables. The degrees of freedom is df, equal to the number of columns in the cross-tabulation minus one, multiplied by the number of rows minus one. A result is significance if the p value is less than or equal to 0.05.

On post-course assessment question 1 (Table 3.4a), Group 3 had higher percentages of students scoring at levels 0, 1, and 3 than Group 1. Group 1 had higher percentages of students at level 2, as well as the top two levels, 4 and 5. For these reasons, Group 1 demonstrated better understanding on question 1 than Group 3.

Table 3.3a: Cross-tabulation for post-course question 3, Groups 1 and 2

Level	Percent within Group 1	Percent within Group 2	Total (n)
0	3%	6%	4
1	49	19	38
2	14	31	19
3	0	9	3
4	0	3	1
5	35	31	33
Total (n)	66	32	98

Table 3.3b: Results for post-course question 3, Groups 1 and 2

	Value	df	p value
Pearson Chi-Square	17	5	0.004

Table 3.4a: Cross-tabulation for post-course question 1, Groups 1 and 3

Level	Percent within Group 1	Percent within Group 3	Total (n)
0	5%	14%	12
1	6	12	12
2	27	17	29
3	38	49	57
4	20	6	17
5	5	2	4
Total (n)	66	65	131

Table 3.4b: Results for post-course question 1, Groups 1 and 3

	Value	df	p value
Pearson Chi-Square	13	5	0.027

Table 3.5a: Cross-tabulation for post-course question 2, Groups 1 and 3

Level	Percent within Group 1	Percent within Group 3	Total (n)
0	12%	12%	16
1	15	23	25
2	15	31	30
3	17	19	23
4	6	5	7
5	35	11	30
Total (n)	66	65	131

Table 3.5b: Results for post-course question 2, Groups 1 and 3

	Value	df	p value
Pearson Chi-Square	13	5	0.023

On post-course assessment question 2, groups 1 and 3 had approximately the same percentages of students scoring at levels 0, 3, and 4, though Group 1 had a slightly higher percentage at level 4 (Table 3.5a). Group 1 also scored much better than Group 3 at level 5. Group 3 had higher percentages at levels 1 and 2 in the rubric. Group 1 outperformed Group 3 at higher rubric levels, surpassing Group 3 in performance on question 2.

Matched Data: Comparing pre- and post-course data within groups

Individual student performances on pre- and post-course assessments for Groups 2 and 3 were compared for all ten questions, using Chi-Square analyses (Table 3.6).

“Was there significant increase in understanding of density from pre- to post-course?”

Cross-tabulations of scores are discussed below for performances on pre-/post-course extended response assessment questions that were significant, e.g., question 1 (Table 3.6) (all non-statistically significant data are in Appendix I).

Table 3.6: Significance of pre-/post-course extended response assessment questions for Groups 2 and 3 (significance at $\alpha = 0.05$)

Question	1	2	3	4	5	6	7	10
Group 2	yes	no	yes	no	yes	yes	yes	no
Group 3	yes	no	yes	yes	no	yes	yes	no

Table 3.7a: Cross-tabulation of pre-/post-course question 1, Group 2

	Post-01						
Pre-01	0	1	2	3	4	5	Total (n)
0	**13%**	38%	13%	25%	13%	0%	8
1	0	**33**	0	67	0	0	3
2	0	13	**76**	13	0	0	8
3	14	0	0	**29**	57	0	7
4	17	0	0	33	**50**	0	6
5	0	0	0	0	0	**0**	0
Total (n)	3	5	7	9	8	0	32

Table 3.7b: Results for pre-/post-course question 1, Group 2

	Value	df	p value
Pearson Chi-Square	31	16	0.014

Group 2 was a small experimental group, with n = 32. Small numbers of students fall at each rubric level: 100% of students may have answered at a certain level, but that may have been only one student. Therefore, discussions of each table concentrate on rubric levels involving more than two students. One student improving or moving down in level is hardly significant. Additionally, the values in the diagonals of each cross-tabulation are bolded: values falling above this bolded diagonal indicate students' improvement in rubric level from pre- to post-course; values falling below the bolded diagonal indicate students' moving down in rubric level from pre- to post-course.

Eighty-seven percent of students within Group 2 who scored a 0 on pre-course assessment question 1 (Table 3.7a) scored at levels 1 to 4 on the post-course assessment, with just over a third of students who first scored a 0 improving to post-course levels 3 or 4. The majority of students who scored a 2 on the pre-course assessment remained at that level on the post-course assessment. Finally, most students who scored 3 or 4 on the pre-course assessment did so on the post-course assessment, and no students scored at level 5.

The following is a sample student response from level 4: "The heating and cooling of magma creates "currents" that shift tectonic plates." Additionally, the student drew a cross-section of the earth with a convection cell. This student's response was typical of the highest level of response for the group: he could discuss convection in terms of temperature differences, but did not cite density differences. The movement to higher rubric levels on the post-course assessment suggests that understanding of mantle convection/plate tectonics did improve from pre- to post-course assessment, but that there was no change in understanding the relation of density to mantle convection/plate tectonics.

Table 3.8a: Cross-tabulation of pre-/post-course question 3, Group 2

	Post-03						
Pre-03	0	1	2	3	4	5	Total (n)
0	**0%**	50%	0%	0%	0%	50%	2
1	18	**18**	46	0	9	9	11
2	0	43	**43**	0	0	14	7
3	0	0	40	**40**	0	20	5
4	0	0	0	0	**0**	0	0
5	0	0	0	14	0	**86**	7
Total (n)	2	6	10	3	1	10	32

Table 3.8b: Results for pre-/post-course question 3, Group 2

	Value	df	p value
Pearson Chi-Square	32	20	0.047

There were no significant shifts to higher rubric levels within Group 2 on question 3 (Table 3.8a). There were small fluctuations in score from pre-course levels 1 to 3 to all rubric levels on the post-course assessment. Notably, most students who first responded at level 5 did so post-course. The following sample response is from level 5: “The fresh water would be above the salt water because the salt water has a higher density, allowing the fresh water to float above it.” There was no dramatic increase in understanding of density in the context of the water well.

Table 3.9a: Cross-tabulation of pre-/post-course question 5, Group 2

	Post-05						
Pre-05	0	1	2	3	4	5	Total (n)
0	**43%**	0%	0%	0%	29%	29%	7
1	0	**100**	0	0	0	0	2
2	0	0	**50**	0	0	50	2
3	33	0	0	**33**	0	33	3
4	0	0	0	0	**67**	33	3
5	0	0	0	0	47	**53**	15
Total (n)	4	2	1	1	11	13	32

Table 3.9b: Results for pre-/post-course question 5, Group 2

	Value	df	p value
Pearson Chi-Square	69	25	0.000

At least half of students within Group 2 scored at the same rubric level on question 5 from pre- to post-course (i.e., levels 0, 4, and 5) (Table 3.9a). Just over half of students who first scored a 0 scored at levels 4 and 5 on the post-course assessment. Improvement in performance at pre-course levels 2, 3, and 4 involved very small numbers of students and so may be misleading, but contributions from these levels resulted in six more students scoring at levels 4 and 5 on the post-course assessment. Finally, all students who scored at levels 4 and 5 on the pre-course assessment did so on the post. The following is a sample student response from level 5: “[The blocks] are all the same material therefore have same density regardless of size.” Overall, aside from fewer students leaving this density question unanswered on the post-course assessment, there was no dramatic change in students’ understanding of density.

Table 3.10a: Cross-tabulation of pre-/post-course question 6, Group 2

	Post-06						
Pre-06	0	1	2	3	4	5	Total (n)
0	**33%**	0%	0%	0%	67%	0%	6
1	50	**5**	0	0	0	0	2
2	0	0	**0**	0	0	100	1
3	50	0	0	**0**	50	0	2
4	6	0	0	0	**81**	13	16
5	0	0	0	0	20	**80**	5
Total (n)	5	1	0	0	19	7	32

Table 3.10b: Results for pre-/post-course question 6, Group 2

	Value	df	p value
Pearson Chi-Square	39	15	0.001

Although students within Group 2 improved on question 6, numbers of students at levels 1, 2, and 3 are small (Table 3.10a). Two-thirds of students who first had no response scored at level 4 on the post-course assessment. The following is a sample student response from level 5: “C is the longest piece, and since all pieces have identical widths and thicknesses, C must have the greatest…mass.” At levels 4 and 5, most responses were identical for pre- and post-course assessment. The results indicate no gain in level of understanding of mass on this question, though score levels were consistently high from pre- to post-course.

Table 3.11a: Cross-tabulation of pre-/post-course question 7, Group 2

	Post-07						
Pre-07	0	1	2	3	4	5	Total (n)
0	**43%**	0%	0%	0%	57%	0%	7
1	0	**0**	0	0	0	0	0
2	0	0	**0**	0	0	0	0
3	0	50	0	**0**	0	50	2
4	0	0	0	8	**39**	54	13
5	10	0	10	0	30	**50**	10
Total (n)	4	1	1	1	12	13	32

Table 3.11b: Results for pre-/post-course question 7, Group 2

	Value	df	p value
Pearson Chi-Square	31	15	0.009

Most improvement within Group 2 on question 7 occurs at level 0: fifty-seven percent of students who first had no response moved up to level 4 on the post-course assessment (Table 3.11a). Most students who first scored at the top two rubric levels did so on the post-course assessment; over half of students who first scored a 4 improved to a score of 5 on the post-course assessment. Half of students originally scoring a 5 dropped in score on the post-course assessment. The following is a sample response from a student who scored a 0 on the pre-course assessment, and scored at post-course level 4: "Density is a measure of mass per volume." By the end of the course, the student demonstrates understanding of the density equation. Overall, more students improved their score level on this question from pre- to post-course, demonstrating improvement in understanding of density.

Table 3.12a: Cross-tabulation of pre-/post-course question 8, Group 2

	Post-08		
Pre-08	a*	b	Total (n)
a*	93%	7%	15
b	35	65	17
Total (n)	20	12	32

Table 3.12b: Results for pre-/post-course question 8, Group 2

	Value	df	p value
Pearson Chi-Square	11	1	0.001

Table 3.13a: Cross-tabulation of pre-/post-course question 9, Group 2

	Post-09		
Pre-09	a*	b	Total (n)
a*	91%	9%	22
b	60	40	10
Total (n)	26	6	32

Table 3.13b: Results for pre-/post-course question 9, Group 2

	Value	df	p value
Pearson Chi-Square	4	1	0.038

Most students within Group 2 remained consistently correct in their choice on question 8 from pre- to post-course (choice a*) (Table 3.12a). Addtionally, 35% of students first choosing incorrectly chose the correct response on the post-course assessment. These results indicate increased students' understanding of density from pre- to post-course.

Slightly more than half of students within Group 2 who first incorrectly chose “b” on question 9 chose the correct response (choice a*) on the post-course assessment (Table 3.13a). These results indicate increased students’ understanding of density from pre- to post-course.

Of ten assessment questions administered to Group 2, four questions (1, 7, 8, and 9) showed significant improvement from pre- to post-course, and three questions (3, 5, and 6) showed significant consistency in response level.

Table 3.14a: Cross-tabulation of pre-/post-course question 1, Group 3

	Post-01						
Pre-01	0	1	2	3	4	5	Total (n)
0	**29%**	17%	17%	38%	0%	0%	24
1	9	**27**	27	36	0	0	11
2	0	0	**25**	75	0	0	8
3	5	5	9	**59**	18	5	22
4	0	0	0	0	**0**	0	0
5	0	0	0	0	0	**0**	0
Total (n)	9	8	11	32	4	1	65

Table 3.14b: Results for pre-/post-course question 1, Group 3

	Value	df	p value
Pearson Chi-Square	25	15	0.046

Seventy percent of students within Group 3 who first scored a 0 on question 1 improved to levels 1 to 3 on the post-course assessment (Table 3.14a). All students who first scored at levels 1 and 2 improved to no higher than level 3 on the post-course assessment. Fifty-nine percent of students who scored a 3 on the pre- course assessment also did so on the post-course assessment. Just over 20% improved from level 3 to levels 4 or 5 on the post-course assessment. Approximately the same number of students fell from pre-course level 3 to levels 0 to 2 on the post-course assessment. Five students scored at levels 4 and 5 on the post-course assessment; no students did so on the pre-course assessment. The following examples demonstrate improvement in one student's response from pre-course level 3 to post-course level 4:

Pre-course: "The molten rock cools and expands causing pressure within the earth's crust causing the plates to move and separate/overlap."

Post-course: "The convection of magma underneath the crust. It churns and moves and the layer touching the lithosphere cools down and attaches itself to it. As the attached magma continues to "churn," it pulls along with it the lithosphere causing movement of the plates."

By the post-course, the student can better explain mantle convection due to temperature differences. Overall, students' understanding of mantle convection/plate tectonics increased from pre- to post-course. However, with only one student scoring at level 5 on the post-course assessment, students' understanding of the role of density in mantle convection/plate tectonics did not drastically change.

Table 3.15a: Cross-tabulation of pre-/post-course question 3, Group 3

	Post-03						
Pre-03	0	1	2	3	4	5	Total (n)
0	**0%**	0%	0%	0%	0%	100%	2
1	9	**77**	3	3	0	9	34
2	0	44	**11**	0	11	33	9
3	17	17	0	**33**	0	33	6
4	0	100	0	0	**0**	0	2
5	0	0	33	17	0	**50**	12
Total (n)	4	33	6	5	1	16	65

Table 3.15b: Results for pre-/post-course question 3, Group 3

	Value	df	p value
Pearson Chi-Square	54	25	0.001

On question 3, seventy-seven percent of students within Group 3 who scored a 1 on the pre-course assessment did so post-course, and 50% of students who scored a 5 on the pre-course assessment did so on the post-course assessment (Table 3.15a). Forty-four percent of students who originally scored a 2 improved to levels 4 or 5 on the post-course assessment, and one-third of students who first scored a 3, improved to level 5 on the post-course assessment. However, 44% of students who first scored a 2 moved down to level 1; two-thirds of students who first scored a 3 moved down to levels 0 to 3; and 50% of students who first scored a 5 moved down to levels 2 and 3. The following is one students' decline in level of response, from pre-course level 5 to post-course level 2:

Pre-course: “The first section would have no water, because it is above sea level. Section two would have fresh water and section three would have salt water because salt water is more dense than fresh water.”

Post-course: [The student predicted the order in the well as: no water above sea level; salt water at sea level; fresh water below sea level.] “When the water trickles through the soil, the salt is precipitated out of it, making it fresh.”

By the end of the course, the student has moved from a scientific explanation involving the relative densities of fresh and salt water, to an (incorrect) idea he has about saltwater filtration. Overall, while many students at the ends of the rubric stayed at the same level from pre- to post-course, changes suggest a decrease in understanding of density in the context of the water well question.

Table 3.16a: Cross-tabulation of pre-/post-course question 4, Group 3

	Post-04						
Pre-04	0	1	2	3	4	5	Total (n)
0	**33%**	0%	0%	0%	33%	33%	3
1	0	**50**	0	0	50	0	2
2	0	0	**0**	0	0	0	0
3	0	0	0	**0**	100	0	2
4	16	0	5	0	**55**	24	38
5	0	0	5	5	15	**75**	20
Total (n)	7	0	4	1	28	25	65

Table 3.16b: Results for pre-/post-course question 4, Group 3

	Value	df	p value
Pearson Chi-Square	31	16	0.013

Though one-third of students within Group 3 who scored at pre-course level 0 on question 4 remained at that level on the post-course assessment, two-thirds improved from pre-course level 0 to levels 4 and 5 on the post-course assessment (Table 3.16a). A few students improved from pre-course assessment levels 1 and 3. Finally, the majority of students who first scored at levels 4 or 5 on the pre-course assessment did so on the post-course assessment, and five more students scored at level 5 than on the pre-course assessment. The following demonstrates one student's improved explanation, from pre-course level 4 to post-course level 5:

Pre-course: "Well, they are all equally dense because they didn't destroy the object, [they] just made it smaller or changed the shape. The large one has a greater volume and mass."

Post-course: "V = LxWxH and the box C has the largest because the length is the greatest." The student's explanation is more scientifically sound by the post-course. Results suggest that student understanding of volume improved from pre- to post-course.

Table 3.17a: Cross-tabulation of pre-/post-course question 6, Group 3

	Post-06						
Pre-06	0	1	2	3	4	5	Total (n)
0	**46%**	18%	0%	0%	18%	18%	11
1	0	**50**	0	17	17	17	6
2	0	0	**0**	0	0	100	1
3	0	2	0	**0**	80	0	5
4	12	6	0	6	**68**	9	34
5	0	0	0	0	63	**38**	8
Total (n)		9	8	3	35	10	65

Table 3.17b: Results for pre-/post-course question 6, Group 3

	Value	df	p value
Pearson Chi-Square	39	20	0.006

Approximately half of students within Group 3 responding at the lowest two levels (0 and 1) on pre-course question 6 remained at those levels on the post-course assessment (Table 3.17a). Thirty-six percent of students who first scored a 0 improved to levels 4 or 5 on the post-course assessment. Fifty percent of students who first scored a 1 improved to levels 3 through 5 on the post-course assessment. Eighty percent of students who first scored a 3 improved to level 4 on the post-course assessment. Finally, the majority of students responding at levels 4 and 5 on the pre-course assessment remained at those levels on the post-course assessment. The following is a sample student response at level 5: "…it's obvious that if compositionally [they are] the same except length, the longest is the heaviest." Results suggest that at the top four levels, students were consistent or improved in understanding of mass from pre- to post-course assessment.

Table 3.18a: Cross-tabulation of pre-/post-course question 7, Group 3

	Post-07						
Pre-07	0	1	2	3	4	5	Total (n)
0	**50%**	0%	0%	0%	0%	50%	4
1	0	**0**	0	0	60	40	5
2	0	0	**0**	0	100	0	1
3	27	9	0	**27**	27	9	11
4	0	4	0	0	**39**	57	23
5	0	5	0	0	19	**76**	21
Total (n)	5	0	3	3	20	34	65

Table 3.18b: Results for pre-/post-course question 7, Group 3

	Value	df	p value
Pearson Chi-Square	47	20	0.001

Improvement within Group 3 on question 7 occurred at all levels except 3 (Table 3.18a). Half of students who first scored at level 0 remained at level 0 on the post-course assessment, whereas 50% of students who first had no response scored at level 5 on the post-course assessment. The following is a sample of one such student's response: "[The density of the diamond chip] is proportional to its m/v, the chip has a smaller mass and volume than the original piece that would equal the same density. It is the same material." By the post-course, this student demonstrated a sophisticated level of understanding of density. Additionally, one hundred percent of students improved from level 1 to levels 4 and 5 on the post-course assessment. Over one-third of students who first scored a 3, improved to levels 4 and 5, while responses of the majority of students at levels 4 or 5 on the pre-course assessment remained at levels 4 and 5 on the post-course assessment. Students showed improvement on understanding of density from pre- to post-course, in the context of this question.

Table 3.19a: Cross-tabulation of pre-/post-course question 8, Group 3

	Post-08		
Pre-08	a*	b	Total (n)
a*	72%	28%	25
b	43	58	40
Total (n)	72	28	65

Table 3.19b: Results for pre-/post-course question 8, Group 3

	Value	df	p value
Pearson Chi-Square	5	1	0.020

Table 3.20a: Cross-tabulation of pre-/post-course question 9, Group 3

	Post-09		
Pre-09	a*	b	Total (n)
a*	72%	28%	36
b	31	69	29
Total (n)	35	30	65

Table 3.20b: Results for pre-/post-course question 9, Group 3

	Value	df	p value
Pearson Chi-Square	11	1	0.001

Half of the students within Group 3 who chose the correct response (choice a*) on question 8 did so on the post-course assessment (Table 3.19a). Additionally, 43% of students who first incorrectly chose “b” chose the correct response on the post-course assessment, demonstrating improvement in students’ understanding of density (albeit modest), from pre- to post-course.

Most students within Group 3 remained consistent in their response on question 9 from pre- to post-course assessment (Table 3.20a). The cross-tabulation reveals no net improvements in understanding of density, from pre- to post-course.

Of ten assessment questions administered to Group 3, five questions (1, 4, 6, 7, and 8) showed significant improvement from pre- to post-course, one question (9) showed consistency in response level, and one question (3) showed significant decline in performance.

Table 3.21: Significance of pre-/post-course multiple-choice assessment questions for Groups 2 and 3 (significance at $\alpha = 0.05$)

Question	2	4	5	6	7	8	9
Group 2	no	yes	yes	yes	no	yes	yes
Group 3	no	yes	yes	yes	yes	yes	yes

We will now examine a few of the results to the multiple-choice portion of the assessment questions (Table 3.21). (Complete multiple-choice question data are in Appendix J.) If we had only relied on multiple-choice questions, our results would have been very different. Recall that Group 2's performances on the extended response portion of question 4 were not statistically significant, from pre- to post-course. Below are the details of the multiple-choice responses to pre- and post-course assessment question 4 for Group 2 (Table 3.22a).

Table 3.22a: Cross-tabulation of pre-/post-course m. c. question 4, Group 2

	Post-04					
Pre-04	a	b	c*	d	e	Total (n)
a	**0%**	0%	0%	100%	0%	1
b	0	**0**	0	0	0	0
c*	4	0	**92**	0	4	28
d	0	0	50	**50**	0	2
e	100	0	0	0	**0**	1
Total (n)	2	0	27	2	1	32

Table 3.22b: Results for pre-/post-course m. c. question 4, Group 2

	Value	df	p value
Pearson Chi-Square	39	9	0.000

Most students in Group 2 chose the correct response (choice c*) to question 4. There were no significant shifts to other choices, and there was no significant improvement in score from pre- to post-course. Because nearly 85% of students chose the correct response on both the pre- and post-course assessments, we may have concluded that most students in the class understood volume, and that there was no change in understanding of volume due to the density laboratory exercise, or the lectures and other laboratory exercises. However, Group 2's pre- and post-course performances on the extended response portion of question 4 were not significantly consistent, as the multiple-choice performances were. In fact, Group 2's variation in performances on the extended response portion was not quite statistically significant, suggesting that students may have been confused about volume. The multiple-choice data combined with the extended response data suggest that the students in Group 2 were able to recognize the correct response, but poorly understood why that response was correct. If only multiple-choice items had been used on the assessment, we would not have known that many students seem to poorly understand the correct answer.

Another example of the need for extended response items is shown by Group 3's performance on assessment question 5. Changes in performance on the extended response portion of the question were not statistically significant, from pre- to post-course, suggesting that students' ability to explain their responses did not change during the course. However, analysis of the multiple-choice portion of the question revealed that 83% of students chose the correct response (choice d) at the end of the course (improvement from only 66% choosing the correct response on the pre-course assessment) (Table 3.23a). If only multiple-choice questions had been employed on the assessment, we may have concluded that student understanding of density (in the context of question 5) improved by the end of the course. It appears that students improved in their ability to recognize the correct response, but not in their ability to explain that correct response.

Inclusion of only extended response items on the assessment may have suggested that students knew less about density than in actuality. The fact that most students can recognize the correct response suggests that they have a partial understanding, but that their understanding is not sophisticated. Certainly, requiring both types of responses demonstrated that a student's ability to choose a correct answer does not necessarily imply complete understanding of the concept.

Table 3.23a: Cross-tabulation of pre-/post-course m. c. question 5, Group 3

	Post-05					
Pre-05	a	b	c	d*	e	Total (n)
a	**40%**	0%	40%	20%	0%	5
b	0	**0**	0	0	0	0
c	0	0	**20**	80	0	5
d*	2	0	0	**93**	5	43
e	0	0	0	75	**25**	12
Total (n)	3	0	3	54	5	65

Table 3.23b: Results for pre-/post-course m. c. question 5, Group 3

	Value	df	p value
Pearson Chi-Square	42	9	0.000

<u>Triangulation within the assessment</u>

Three other questions are addressed:

- Were students' responses consistent on questions 5 and 7?
- Were students' responses consistent on questions 8 and 9?
- Were students' responses consistent on questions 4, 5, and 6?

Results of performances on the triangulation assessment questions are summarized below in Table 3.24. Cross-tabulations of scores are discussed for statistically significant performances (all non-statistically significant data are in Appendix J).

Table 3.24: Significance of pre-/post-course assessment triangulation questions for Groups 2 and 3 (significance at $\alpha = 0.05$)

Questions	5 and 7		8 and 9		4 and 5		4 and 6		5 and 6	
	pre	post	pre	post	pre	post	pre	post	pre	post
Group 2	yes	no	no	no	yes	yes	yes	no	yes	yes
Group 3	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

"Were students' responses consistent on questions 5 and 7?"

The largest percentages of students' scores within Group 2 on pre-course assessment questions 5 and 7 occur at either end of the rubric and demonstrate consistency in responses: most students who scored at level 0 on question 5 did so on question 7; 100% of students who scored at level 5 on question 5 scored at the top two levels on question 7 (Table 3.25a).

There was some inconsistency in level of score within Group 3 on pre-course assessment questions 5 and 7: over half of the students who first scored a 0 on question 5 scored at levels 2 to 5 on question 7; seventy-five percent of students who first scored a 1 scored at the top three levels on question 7; nearly 70% of students who first scored at level 3 scored at levels 4 and 5 on question 7; and the majority of students who first responded at levels 4 and 5 did so on question 7 (Table 3.26a). For question 5 to 7, responses were either consistent, or better on question 7.

Table 3.25a: Cross-tabulation of pre-course questions 5 and 7, Group 2

	Pre-07						
Pre-05	0	1	2	3	4	5	Total (n)
0	**71%**	0%	0%	0%	14%	14%	7
1	0	**0**	0	50	0	50	2
2	0	0	**0**	0	50	50	2
3	33	0	0	**33**	33	0	3
4	33	0	0	0	**0**	67	3
5	0	0	0	0	67	**33**	15
Total (n)	7	0	0	2	13	10	32

Table 3.25b: Results for pre-course questions 5 and 7, Group 2

	Value	df	p value
Pearson Chi-Square	33	15	0.005

Table 3.26a: Cross-tabulation of pre-course questions 5 and 7, Group 3

	Pre-07						
Pre-05	0	1	2	3	4	5	Total (n)
0	**43%**	0%	14%	14%	14%	14%	7
1	0	**25**	0	38	31	6	16
2	0	0	**0**	100	0	0	1
3	0	8	0	**23**	31	39	13
4	5	0	0	0	**55**	41	22
5	0	0	0	0	17	**83**	6
Total (n)	4	5	1	11	23	21	65

Table 3.26b: Results for pre-course questions 5 and 7, Group 3

	Value	df	p value
Pearson Chi-Square	62	25	0.000

On the post-course assessment, seventy-five percent of students within Group 3 who scored at a level 0 on question 5 scored at levels 4 and 5 on question 7 (Table 3.27a). Over 60% of the students who scored at level 3 on question 5 scored at levels 4 and 5 on question 7. Additionally, the majority of students who scored at levels 4 and 5 on question 5 did so on question 7. These results indicate that students' level of response improved from questions 5 to 7 on the post-course assessment.

Table 3.27a: Cross-tabulation of post-course questions 5 and 7, Group 3

	Post-07						
Post-05	0	1	2	3	4	5	Total (n)
0	**25%**	0%	0%	0%	25%	50%	4
1	29	**0**	0	14	57	0	7
2	0	100	**0**	0	0	0	1
3	9	9	0	**18**	18	46	11
4	3	0	0	0	**35**	62	34
5	0	13	0	0	13	**75**	8
Total (n)	5	3	0	3	20	34	65

Table 3.27b: Results for post-course questions 5 and 7, Group 3

	Value	df	p value
Pearson Chi-Square	48	20	0.000

"Were students' responses consistent on questions 8 and 9?"

On the pre-course assessment, the majority of students within Group 3 who chose "b" on question 8 did so for question 9 (Table 3.28a). Similarly, the majority of students who correctly chose "a" on question 8 did so on question 9. Students' demonstrated level of understanding of density was consistent from question 8 to question 9.

On the post-course assessment, most students within Group 3 were consistent in their choice from question 8 to question 9 (Table 3.29a). Students' demonstrated level of understanding of density was consistent from question 8 to question 9. Additionally, from pre- to post-course, there was a small increase (6 students) in the number of students scoring correctly on both questions.

Table 3.28a: Cross-tabulation of pre-course questions 8 and 9, Group 3

	Pre-09		
Pre-08	a*	b	Total (n)
a*	84%	16%	25
b	38	63	40
Total (n)	36	29	65

Table 3.28b: Results for pre-course questions 8 and 9, Group 3

	Value	df	p value
Pearson Chi-Square	13	1	0.000

Table 3.29a: Cross-tabulation of post-course questions 8 and 9, Group 3

	Post-09		
Post-08	a*	b	Total (n)
a*	77%	23%	35
b	27	73	30
Total (n)	35	30	65

Table 3.29b: Results for post-course questions 8 and 9, Group 3

	Value	df	p value
Pearson Chi-Square	17	1	0.000

"Were students' responses consistent on questions 4, 5 and 6?"

A cross-tabulation can only compare two sets of data. To compare questions 4, 5, and 6, we used three separate Chi-squared analyses: (1) questions 4/5; (2) 4/6; and (3); 5/6.

Table 3.30a: Cross-tabulation of pre-course questions 4 and 5, Group 2

	Pre-05						
Pre-04	0	1	2	3	4	5	Total (n)
0	**80%**	0%	20%	0%	0%	0%	5
1	0	**0**	0	0	0	0	0
2	0	0	**0**	0	0	0	0
3	0	0	0	**67**	0	33	3
4	18	18	9	0	**9**	46	11
5	8	0	0	8	15	**69**	13
Total (n)	7	2	2	3	3	15	32

Table 3.30b: Results for pre-course questions 4 and 5, Group 2

	Value	df	p value
Pearson Chi-Square	33	15	0.004

Table 3.31a: Cross-tabulation of pre-course questions 4 and 6, Group 2

	Pre-06						
Pre-04	0	1	2	3	4	5	Total (n)
0	**80%**	0%	20%	0%	0%	0%	5
1	0	**0**	0	0	0	0	0
2	0	0	**0**	0	0	0	0
3	0	0	0	**33**	67	0	3
4	9	9	0	0	**82**	0	11
5	8	8	0	8	39	**39**	13
Total (n)	6	2	1	2	16	5	32

Table 3.31b: Results for pre-course questions 4 and 6, Group 2

	Value	df	p value
Pearson Chi-Square	35	15	0.002

Table 3.32a: Cross-tabulation of pre-course questions 5 and 6, Group 2

	Pre-06						
Pre-05	0	1	2	3	4	5	Total (n)
0	**86%**	0%	0%	0%	14%	0%	7
1	0	**50**	0	0	50	0	2
2	0	0	**50**	0	50	0	2
3	0	0	0	**33**	67	0	3
4	0	33	0	0	**67**	0	3
5	0	0	0	7	60	**33**	15
Total (n)	6	2	1	2	16	5	32

Table 3.32b: Results for pre-course questions 5 and 6, Group 2

	Value	df	p value
Pearson Chi-Square	60	25	0.000

Most students within Group 2 scored at the same level of response for questions 4 and 5 (Table 3.30a), except level 4 question 4, where approximately half of the students dropped to levels 0 to 3, and approximately half of the students scored at levels 4 and 5.

Most students within Group 2 responded consistently to questions 4 and 6 (Table 3.31a). An exception is question 4: the majority of students who responded at level 5 answered at level 4 or 5 on question 6. The majority of students had the same level of response for questions 5 and 6 (Table 3.32a). For questions 4, 5, and 6 on the pre-course assessment, level of understanding was consistent for the majority of students within Group 2.

Table 3.33a: Cross-tabulation of post-course questions 4 and 5, Group 2

	Post-05						
Post-04	0	1	2	3	4	5	Total (n)
0	**50%**	0%	0%	0%	50%	0%	4
1	50	**0**	0	0	50	0	2
2	0	0	**0**	0	0	0	0
3	0	0	0	**50**	50	0	2
4	0	20	10	0	**20**	50	10
5	7	0	0	0	36	**57**	14
Total (n)	4	2	1	1	11	13	32

Table 3.33b: Results for post-course questions 4 and 5, Group 2

	Value	df	p value
Pearson Chi-Square	36	20	0.016

Table 3.34a: Cross-tabulation of post-course questions 5 and 6, Group 2

	Post-06						
Post-05	0	1	2	3	4	5	Total (n)
0	**75%**	0%	0%	0%	25%	0%	4
1	0	**50**	0	0	50	0	2
2	0	0	**0**	0	100	0	1
3	0	0	0	**0**	100	0	1
4	18	0	0	0	**73**	9	11
5	0	0	0	0	54	**46**	13
Total (n)	5	1	0	0	19	7	32

Table 3.34b: Results for post-course questions 5 and 6, Group 2

	Value	df	p value
Pearson Chi-Square	35	15	0.003

On the post-course assessment, students within Group 2 were generally consistent in their responses to questions 4 and 5. Most students who scored at level 0 or 5 on question 4 did so on question 5 (Table 3.33a).

On the post-course assessment, students within Group 2 were generally consistent in their responses to questions 5 and 6 (Table 3.34a).

Table 3.35a: Cross-tabulation of pre-course questions 4 and 5, Group 3

	Pre-05						
Pre-04	0	1	2	3	4	5	Total (n)
0	**67%**	0%	0%	33%	0%	0%	3
1	0	**50**	50	0	0	0	2
2	0	0	**0**	0	0	0	0
3	50	0	0	**50**	0	0	2
4	9	32	0	18	**37**	5	38
5	5	15	0	20	40	**20**	20
Total (n)	7	16	1	13	22	6	65

Table 3.35b: Results for pre-course questions 4 and 5, Group 3

	Value	df	p value
Pearson Chi-Square	56	20	0.000

Table 3.36a: Cross-tabulation of pre-course questions 4 and 6, Group 3

	Pre-06						
Pre-04	0	1	2	3	4	5	Total (n)
0	**67%**	0%	0%	33%	0%	0%	3
1	0	**50**	0	50	0	0	2
2	0	0	**0**	0	0	0	0
3	0	0	0	**0**	50	50	2
4	13	8	3	8	**66**	3	38
5	20	10	0	0	40	**30**	20
Total (n)	11	6	1	5	34	8	65

Table 3.36b: Results for pre-course questions 4 and 6, Group 3

	Value	df	p value
Pearson Chi-Square	35	20	0.022

On the pre-course assessment, most students within Group 3 remained at each rubric level from questions 4 to 5 or improved (Table 3.35a). Sixty percent of students who scored a 5 on question 4 remained at levels 4 or 5 on question 5. Two-thirds of students who scored a 0 on question 4 did so on question 5. Over half of the students who first responded at level 4 on question 4 scored at levels 0 to 3 on question 5.

On the pre-course assessment, most students within Group 3 remained at each rubric level from question 4 to 6 (Table 3.36a). For example, approximately 70% of students who first responded at level 4 scored at levels 4 and 5 on question 6. Seventy percent of students who first responded at level 5 scored at levels 4 and 5 on question 6. However, nearly 30% of students, who first responded at level 4, scored at levels 0 to 3 on question 6. Additionally, 30% of students who first responded at level 5 scored at levels 0 and 1 on question 6.

On the pre-course assessment, most students within Group 3 remained at each rubric level from questions 5 to 6 or improved (Table 3.37a). Most students who responded at levels 4 and 5 on question 5 did so on question 6. Slightly fewer than half of the students who first scored at level 0 responded at level 4 on question 6. Approximately 70% of students who scored a 1 on question 5 improved to levels 3 to 5 on question 6. Nearly half of students who scored at level 3 on question 5 responded at level 4 on question 6.

Table 3.37a: Cross-tabulation of pre-course questions 5 and 6, Group 3

	Pre-06						
Pre-05	0	1	2	3	4	5	Total (n)
0	**57%**	0%	0%	0%	43%	0%	7
1	6	**25**	0	13	50	6	16
2	0	0	**0**	100	0	0	1
3	15	15	0	**15**	46	8	13
4	14	0	0	0	**73**	14	22
5	17	0	17	0	17	**50**	6
Total (n)	11	6	1	5	34	8	65

Table 3.37b: Results for pre-course questions 5 and 6, Group 3

	Value	df	p value
Pearson Chi-Square	54	25	0.001

Table 3.38a: Cross-tabulation of post-course questions 4 and 5, Group 3

	Post-05						
Post-04	0	1	2	3	4	5	Total (n)
0	**43%**	14%	0%	14%	29%	0%	7
1	0	**50**	0	25	25	0	4
2	0	0	**0**	0	0	0	0
3	0	0	0	**100**	0	0	1
4	4	14	4	18	**50**	11	28
5	0	0	0	12	68	**20**	25
Total (n)	4	7	1	11	34	8	65

Table 3.38b: Results for post-course questions 4 and 5, Group 3

	Value	df	p value
Pearson Chi-Square	38	20	0.009

Table 3.39a: Cross-tabulation of post-course questions 4 and 6, Group 3

	Post-06						
Post-04	0	1	2	3	4	5	Total (n)
0	**57%**	14%	0%	0%	29%	0%	7
1	0	**75**	0	0	25	0	4
2	0	0	**0**	0	0	0	0
3	0	0	0	**0**	100	0	1
4	7	7	0	7	**75**	4	28
5	12	8	0	4	40	**36**	25
Total (n)	9	8	0	3	35	10	65

Table 3.39b: Results for post-course questions 4 and 6, Group 3

	Value	df	p value
Pearson Chi-Square	42	16	0.000

Table 3.40a: Cross-tabulation of post-course questions 5 and 6, Group 3

	Post-06						
Post-05	0	1	2	3	4	5	Total (n)
0	**100%**	0%	0%	0%	0%	0%	4
1	0	**71**	0	0	29	0	7
2	0	0	**0**	100	0	0	1
3	9	9	0	**0**	73	9	11
4	8	3	0	3	**65**	21	34
5	13	13	0	13	38	**25**	8
Total (n)	9	8	0	3	35	10	65

Table 3.40b: Results for post-course questions 5 and 6, Group 3

	Value	df	p value
Pearson Chi-Square	77	20	0.000

On the post-course assessment, most students within Group 3 performed consistently or scored higher on question 5 than on question 4 (Table 3.38a). Fifty-seven percent of students who scored a 0 on question 4 improved to levels 1 to 5 on question 5. Half of students who scored a 1 on question 4 improved to levels 3 to 4 on question 5. Additionally, most students who scored at levels 4 and 5 on question 4 remained so on question 5.

On the post-course assessment, most students within Group 3 performed consistently on questions 4 and 6, except that most students who scored at level 5 on question 4 scored at levels 4 and 5 on question 6 (Table 3.39a).

On the post-course assessment, many students within Group 3 scored higher on question 6 than on question 5 (Table 3.40a). Eighty-two percent of students who scored a 3 on question 5 improved to levels 4 and 5 on question 6. Additionally, most students who scored at level 4 or 5 on question 5 did so on question 6. Overall, students' responses were consistent, and commonly improved, from question 4 to questions 5 and 6.

Of the seven triangulation questions administered to Group 2, performances on questions 5 and 7 on the pre-course assessment were consistent, and performances on questions 4, 5, and 6 were generally consistent, both pre- and post-course.

Of the seven triangulation questions administered to Group 3, most students performed consistently or scored higher on question 7 than on question 5, both pre- and post-course; performances were consistent on questions 8 and 9; and performances on questions 4, 5, and 6 were consistent or improved from question to question, both pre- and post-course.

Discussion

Population Data: Comparing post-course data between groups

Groups 2 and 3 were compared with Group 1 in a series of Chi-squared analyses. Each analysis dealt with responses to one of three questions about density in an earth science context, with all three questions making up the first portion of the pre-/post-course assessment.

The Chi-squared analyses comparing Groups 1 and 2 on the three contextual questions determined that performances on only question 3 of the three questions were statistically different among groups. Question 3 dealt with a well, drilled parallel to an ocean-side cliff. Students were asked to predict the order in which no water (air), fresh water, and salt water would be found within the well. Group 2 slightly outperformed Group 1 at levels 2 to 4 on the rubric, and both groups had similar performance at level 5. There are several possible reasons why Group 2 outperformed Group 1 on question 3. Perhaps the professor teaching Group 2 discussed the physical properties of salt versus fresh waters more than the professor teaching Group 1. I sat in on all lectures given to Group 1, but only a few given to Group 2, and so I cannot rule out the possibility of professor-influence. Or, perhaps some students in Group 2 had dealt previously with the density of water. Pre-/post-course assessment performances on question 3 for Group 2 showed no real change. This fact suggests that there was no learning in the course related to question 3, either from lecture or laboratory exercises. In fact, the laboratory exercise on density did not explicitly address the density of air or salt water, and so may not have influenced student understanding on question 3.

There were no real differences in performance between Groups 1 and 2 for contextual questions 1 and 2. We hypothesized that Group 2 would outperform Group 1, because of the specially designed density laboratory exercise. The density laboratory exercise did not significantly influence Group 2's understanding of density in an earth science context.

Chi-squared analyses of Groups 1 and 3 were no more encouraging as to the usefulness of the density laboratory exercise. The cross-tabulations for questions 1 and 2 show that Group 1 outperformed Group 3 rubric levels 4 and 5. Groups 1 and 3 performed equally on question 3. These results suggest that Group 1 had a better understanding of density as it relates to plate tectonics, including mantle convection and subduction mechanisms. The density laboratory exercise did not explicitly include these topics. The same professor taught Groups 1 and 3, one year apart. There was no more emphasis on plate tectonics one year compared to the next (personal observations and course syllabi) and much information about plate tectonics came from reading assigned in the course, using the same textbook for both years. These facts suggest that students in Group 1 initially had a better understanding of the role of density in mantle convection/plate tectonics than students in Group 3. Perhaps more of those students had taken physics or chemistry courses in which density was discussed. Or, perhaps Group 1 had a higher proportion of students who had been exposed to the theory of plate tectonics and the concept of density at the primary or secondary levels, while fewer students in Group 3 had encountered such concepts.

Answers to such conjectures remain speculation, because we do not have pre-course data on Group 1. We do not know what concepts or prior knowledge the students had upon entering the class, and so we do not know whether the course influenced their levels of understanding of density. Furthermore, we do not know any of the students' backgrounds, or what other classes they were taking concurrently with ERS 102, which may have played a role in influencing their levels of understanding of density. Therefore, we cannot use results from Group 1 (post-course) to draw any meaningful conclusions about the usefulness of the density laboratory. Group 1 allowed us to develop and test some of the questions on our assessment, but strict comparisons with subsequent groups is not valid.

Matched Data: Comparing pre- and post-course data within groups

1. Group 2

Of ten assessment questions administered to Group 2, four questions (1, 7, 8, and 9) showed significant improvement in response level from pre- to post-course, and three questions (3, 5, and 6) showed consistency in response level. Additionally, three questions (2, 4, and 10) showed statistically insignificant differences in response level.

The density laboratory exercise did not explicitly deal with density in the context of mantle convection or plate tectonics, which were the topics of questions 1 and 2, respectively. Improvement in level of understanding of these concepts most likely came from the lecture or the textbook. Students were assigned reading on plate tectonics, and studied related rock groups and rock formation environments in laboratory. These factors, combined with any other outside influences, such as specials on television or other courses taken concurrently with ERS 102, may have helped to boost students' scores on the first assessment question, but did not do so on question 2. In fact, question

2 responses were inconsistent from pre-/post-course assessment. Perhaps this multiple-choice question was too confusing or tricky for students to answer correctly.

The lack of improvement on question 3 is not surprising, because the laboratory exercise on density never explicitly explored the density differences among air and fresh and salt water. The use of water displacement to determine densities, and the suggestion that students determine the density of the water they used in the exercise might have helped. Additionally, there was an extra-credit question asking about the density of pumice, which students could observe as being filled with air pockets. The fact that the rock floated on water could have helped students understand that air is less dense than water, but it was never explicitly mentioned in the exercise or discussed in class.

Questions 4, 5, and 6 were all interrelated, dealing with a uniform board cut into three differently sized pieces. Changes in student performance on the volume question, 4, were not statistically significant, though scores were high both pre- and post-course. Perhaps some students remained unsure of how to measure the volume of a rectangular solid, or had forgotten the appropriate equation, post-course. The laboratory exercise dealt with finding the volume of an irregularly shaped object, and the volume equation for rectangular solids was not used. Alternatively, perhaps students retained volume misconceptions from pre- to post-course. Performances on questions 5 and 6, asking about the density and mass of the objects, respectively, were consistent but showed no improvement from pre- to post-course. Some students cited a need for measurements before they could answer the density question. Lack of improvement in scores suggests that students who were confused about mass and density remained so at the end of the course, however most students demonstrated an understanding of those concepts.

Improvement on questions 7, 8, and 9 is exciting, because those questions address the concept of density out of an earth science context (though the topic of question 7 is a diamond chip, the question falls in Part II of the assessment). All three questions ask about the density of one object compared to another, which is a topic the laboratory exercise on density did explicitly cover. Furthermore, question 7 deals with a homogenous substance, a diamond; the density exercise deals with homogeneous/heterogeneous substances and their respective densities. Improvement on question 7 optimistically may be attributed partly to the laboratory exercise, as well as students learning in the course that diamonds are homogeneous. Any student who had never previously thought about the make-up of a diamond may have been confused on the pre-course assessment, but may have had a better understanding on the post-course assessment. Nearly every student who did not score correctly on the pre-course assessment shifted to correct on the post-course assessment.

Both of questions 8 and 9 ask about density, but in different ways. Question 8 asks about density as a verbalization of the density equation: “which object has a greater mass for its volume?” Question 9 asks, in reference to the same two objects as in question 8: “which object has a greater density?” The density laboratory exercise dealt with all of the concepts in questions 8 and 9. Students’ performances improved on both of these questions, even though we regarded question 8 as more difficult, because wording was not as straightforward as in question 9. Also, of the students who got question 9 correct, only roughly one-half got question 8 correct (Appendix I, Table I.12a). These results suggest that students in Group 2 had a clear understanding the way in which mass and volume relate to density.

There were no significant differences in scores from pre-/post-course assessment on question 10. This is most likely because the question was too confusing, and possibly too difficult, for most students. The question was confusing because the same variables, A and B, were used to designate both the blocks in question 10 and in questions 8 and 9. However, in 8 and 9, the blocks were balanced, clearly having the same mass. In question 10, the blocks were unbalanced, clearly having different masses, and so could not have been the same blocks as in 8 and 9. Different letters should have been used in question 10 to avoid confusion among questions. The wording of question 10 suggested that one or the other block should have a greater density; answering "neither" was not given as an alternative, even though it was the correct answer. For this reason, students may have felt that their intuition to write "neither" was wrong, and simply chose a block.

Stavy and Tirosh (2000) may have a different interpretation for students who appeared to be confused by the blocks. When the "intuitive rule," that "more A = more B," is applied to question 10, students may have thought that object B was denser because it tipped the scale. In fact, many students had just that explanation for choosing block B: "Because B is heavier, density is greater, " and, "more density creates a 'heavier' solid." These responses were categorized at the developed/superficial level of understanding, because a student's explanation appeared to incorporate density, but understanding was superficial and incorrect. In short, the students knew that they should talk about density, and they tried to do so, but fell short. The percent of responses at rubric level 3 increased from pre- to post-course, for both groups. Perhaps even after completing the laboratory exercise on density, some students believed that more weight means a greater density for an object. More questions are necessary on the assessment, in order to separately question students on weight and density, and determine whether those two concepts were confused.

Question 10 may have been difficult because the students did not know the mass-to-volume ratio for each block (density); many students cited a need for specific measurements, as they had used in the laboratory exercise. Students with a strong grasp of the density equation (D = M/V) could conclude which block had the greater density, provided that they were confident enough to write “neither,” and not second-guess themselves. These factors likely diminish our ability to determine a student’s true level of understanding.

2. Group 3

Of ten assessment questions administered to Group 3, five questions (1, 4, 6, 7, and 8) showed significant improvement in performance level from pre- to post-course, one question (9) showed consistency in response level, and one question (3) showed significant decline in performance. Additionally, three questions (2, 5, and 10) yielded statistically insignificant results.

Group 3 had more questions showing significant improvement than Group 2. Performances on questions 1, 2, 3, 7, and 10 from pre- to post-course may have been influenced by factors suggested for Group 2.

Questions 5, 7, 8, 9, and 10 all dealt explicitly with density. Perhaps the two questions showing improvement in scores from pre- to post-course (questions 7 and 8) were easier for students to understand. However, improvement on question 8 and only consistent performance on question 9 is surprising (Table 3.28a and 3.29a). We expected most students to perform better on question 9 than on question 8, because question 9 asks about density in a more straightforward manner. The fact that three of six density questions did not show significant improvement in scores from pre- to post-course suggests that some students may have remained confused about density.

There was improvement in pre- to post-course scores on questions 4 and 6, dealing with the volume and mass of three boards, respectively. Students seemed to improve their understanding of the components of the density equation, but could not show similar improvement on a density question (number 5) in the same context as the other two questions. This may be due to the order of the questions: volume, density, and mass. Perhaps moving from thinking about volume to thinking about density, instead of thinking about mass and then thinking about density, as verbalization of the density equation may suggest students do, was confusing for those with a weak grasp of the concept. Alternatively, perhaps the assessment we developed for this study was not sensitive enough to detect slight changes in understanding of density, mass, or volume (though slight changes in understanding may not be very meaningful).

Triangulation within the assessment

1. Group 2

Part II of the assessment consists of density questions, using different materials and various situations. Repetition was intentionally built into the assessment, to better discern a student's level of understanding. The repetition of questions, i.e., triangulation, is a necessary part of any reliable education research instrument. The first pair of questions analyzed together was 5 and 7; both questions asked about the density of an object. For Group 2, there was some consistency in pre-course response from question 5 to 7, but post-course responses to questions 5 and 7 were inconsistent.

Questions 8 and 9 also attempted to triangulate students' understanding of density. Group 2's performance on question 8 was inconsistent with performance on question 9. As previously discussed, students may have been confused by the wording in question 8, which asked the students to name the block with "greater mass for its volume." Any student with a solid grasp of the concept of and the equation for density should have had no problem with questions 8 and 9, unless the wording was simply too confusing. This suggests that some students only had a superficial understanding of density at the end of the course.

Questions 4, 5 and 6 asked about volume, density, and mass, respectively. Group 2 performed consistently from one question to the next on the pre-course assessment, but did not do so on the post-course assessment. Students who scored at one level on question 4 did not necessarily do so on question 6. This suggests that at the end of the course, students in Group 2 could understand isolated portions of the density equation: volume in relation to density; and density in relation to mass. But many students appeared to have trouble understanding mass and volume at the end of the course, and perhaps were simply guessing on question 6. More mass and volume assessment questions could have allowed us to better gauge students' levels of understanding of mass and volume.

2. Group 3

Pre-course assessment performances on questions 5 and 7 were consistent for Group 3, and post-course assessment scores were higher on question 7 than on question 5. Perhaps question 5 was more confusing to students than question 7. This idea is supported by pre-/post-course assessment comparisons: there was improvement on responses to question 7, but changes in score on question 5 from pre- to post-course were statistically insignificant. Alternatively, these results may suggest that students' understanding of density was incomplete by the end of the course. We expected that students who had a sophisticated understanding of density would perform consistently at high levels on questions 5 and 7.

Pre-/post-course assessment performances of questions 8 and 9 for Group 3 were consistent from question to question. However, students actually performed slightly worse on question 9 on the post-course assessment. Despite the consistency from question 8 to 9, students in Group 3 may have been more confused about density at the end of the course.

Pre-/post-course assessment performances on questions 4 through 6 for Group 3 were consistent from one question to the next. Students could move from the volume question, to the density question, and finally to the mass question, generally without any prompting from the questions or learning from question to question. Consistency in level of understanding suggests that there was very little guessing on the part of most students. Finally, consistency in score from questions 4 to 6 suggests that the scoring of responses was reliable. Reliability in grading was a concern for this study, because of the subjective nature of scoring the responses; human error will always be involved in education research.

Limitations of the study design

Individual performances within the baseline group were not tracked from pre- to post-course, as mentioned above. The data collected from the baseline group (Group 1) were only post-course population data. The baseline data were compared to population data from experimental groups (Groups 2 and 3), and we cannot be sure that factors other than the intervention were the cause of variation among groups. Furthermore, there were no control groups for the statistical comparisons made for Groups 2 and 3. Any improvement in scores on post-course assessments is attributed to the intervention, though that may not be the case. For example, students learning about density in classes taken concurrently with ERS 102 may have performed better on the post-course assessment, regardless of any benefit from having completed the density laboratory exercise.

Limitations of the instruments

1. The density laboratory exercise

The exercise was lacking in that it:

- never explicitly addressed convection in the mantle or in fluids;
- may have been too far removed from the context of earth sciences (though part of the point of this study was to determine whether a basic “primer” in density would help students in earth science);

- explored only hetero-/homogeneous substances at constant temperature and pressure, and did not explore other causes of variation of density;
- employed the water displacement method, though we wished to avoid buoyancy-related phenomena;
- never asked students to make a clear distinction between mass and weight, or to use only mass (not weight) in the density equation in laboratory reports. This may have confused students, because they have probably learned in other science classes that the weight of an object can vary in different environments. Driver et al. (1985) noted that the concept of mass is confusing for students, possibly because weight is commonly misused in place of mass. Students seem to have the idea that mass can vary for an object, whereas weight is what varies in different environments. Failure to make a distinction between mass and weight in the laboratory exercise may have led to confusion about density varying in certain situations, when the students should have been learning that density is a material property, and does not vary for a homogenous substance.
- was never linked to subsequent laboratory exercises; and it
- never linked to any homework/thought questions to keep density-related ideas alive outside of the classroom (aside from examination questions).

The exercise we designed was allotted one laboratory period (2 hours), and students were rushed to finish with only the two parts that had been included in the exercise; many students never got to determine the density of their rock/mineral repeatedly, to obtain a more accurate value, and many groups did not reach the extra credit question.

Finally, assessment results from Group 2, the first experimental group, should have been used to improve the laboratory exercise. For Group 2, questions 3, 5, 6, and 8 yielded no change in score level from pre- to post-course; these results could have been used to revise the laboratory exercise to target those areas more closely. Statistical analyses for both groups were done at the same time, too late for revisions based on the performance of the earlier group's performance.

2. The assessment

The assessment designed for this study was very limited in terms of the quality of questions employed and the number of revisions. The three contextual questions in Part I (Appendix C) were developed by members of this research project over several iterations, and were believed to capture the concept of density in contexts the students were likely to encounter in ERS 102. Additionally, various graduate students and Group 1, the baseline group, tested Part I of the assessment before it was implemented for the experimental groups. Finally, questions 4 to 7 of Part II of the assessment were taken from a survey, which had already undergone many revisions (Yeend et al., 2002; Loverude, 2004). We had confidence that these questions would yield useful data. By keeping assessment questions straightforward, and by requiring the proper explanation to include density, any omission of density would suggest a partial or imperfect understanding of that concept. However, that includes the assumption that each question would accurately elicit and gauge a student's level of understanding of density.

Based upon earlier results, including results from Group 2, the questions that returned non-statistically significant results (questions 2, 4, and 10) should have been revised for clarity, content, or perhaps removed from the instrument altogether. A new iteration of the assessment should have been conducted with Group 3. The timeline for this project did not enable iterations, especially with only two experimental groups. Additionally, more introductory earth science classes should have been asked to participate in the study, in order to increase the size of our data set, as well as to diversify the science background in our study participants.

Limitations of the analysis

The rubrics introduced more error into the study. Each level was assigned a set of key ideas to be used as guidelines when assigning a student's response to a certain level of understanding. The rubrics were modified from examples in Wiggins and McTighe (1998). They hint at the error inherent in scoring students' responses: "Clearly, understanding is a matter of degree on a continuum. It is not a matter of right versus wrong but *more* or *less* naïve or sophisticated; *more* or *less* superficial or in-depth… Whatever the response, rubrics provide useful guidance in assessment." Many students' responses fell into the "gray areas" between two rubric levels, and the key ideas at those levels were adjusted to accommodate the responses.

Many students' remarks were very glib, leading to difficulty in scoring. These students may have assumed that we would be familiar with their level of understanding, after having spent a semester working with them in laboratory and grading their exercises, and perhaps they believed a complete explanation was not needed. We did not give any student the benefit of the doubt. For this study, we were limited to the responses on the assessments. Many students left fragmented or incomplete responses, and they were graded accordingly. In this way, we tried to approach every student in the same way: what level of understanding could they demonstrate, given the assessment questions?

Also, students had only one space for the explanation of questions 4, 5, and 6. Students commonly explained their choices for questions 4 to 6 in one long sentence. For this reason, a student's explanation for question 4 may have been used to help classify his explanation to questions 5 or 6.

The scoring of students' responses at the superficial level of understanding (level 3) introduced more error into the study. Most responses that fell into this level contained technical language, but no substance to back up the vocabulary. At the end of the course, many students were proficient in technical language for earth science processes, or could name a certain process, whereas in the beginning of the course, very few students used technical language in responding to the pre-course assessment questions. Though questions explicitly asked students to explain their position, perhaps students assumed that mentioning a technical earth science word was sufficient explanation. We did not give these students the benefit of the doubt; we assumed that they only knew the terminology, because they had not demonstrated the extent of their understanding of that term. The problem of superficial learning, shown by the increased use of scientific terminology with no ability to explain those terms, is not a new one. Stepans et al. (1986) found that after his students completed a laboratory exercise on sinking and floating, "unfortunately, the junior high students' increase in sophisticated science vocabulary was not accompanied by increased understanding."

For our study, perhaps a word bank should have been available on the assessment. Each question would ask the students to pick the appropriate word to use in their explanation, and the question would ask the student to first define the word. In that way, we could more accurately assess the level of understanding that the student possessed, and have a better gauge of a student's mastery of a subject.

Limitations of the research participants

Apathy on the part of a student is the bane of every teacher's professional existence. Only a few students opted out of the study, and participating had no immediate benefits. The assessments were not graded, there was no extra credit allotted for participation, and no candy bars were given out as bribes! The students who chose to participate in the study were self-selected, and probably attempted to do their best on the assessment questions.

Confidence may have been the biggest factor in whether a student chose to respond to a question. In both groups, more students attempted to answer questions on the post-course assessment than on the pre-course assessment (i.e., there were fewer scores of 0 on the post-course assessments, overall). For example, 5 more students in Group 2 attempted to respond to question 1 on the post-course assessment than had done so on the pre-course assessment. For Group 3, 15 more students attempted to respond to question 1 on the post-course assessment than had done so on the pre-course assessment. Some students may still have been confused about density at the end of the course, or did not want to respond to additional difficult questions. Our data suggest that students attempted to respond to questions if they could.

Conclusions

Summary comments include:

(1) Analyses of performances on pre-/post-assessment questions revealed one-third to one-half of questions with significant improvement in understanding of density for both experimental groups. This suggests that the inquiry-based laboratory exercise may have been useful in improving student understanding of density. However, lack of pre-course data from Group 1 (the baseline group) precludes a firm conclusion about the true efficacy of the exercise.

(2) Use of multiple-choice questions may be misleading in terms of uncovering the extent of student understanding of a concept. Use of multiple-choice and extended response items on our assessments allowed for a broader look at student understanding of density.

(3) Revision of the 2004 version of the density laboratory exercise, to a more inquiry-based exercise for the spring 2005 group, did not seem to have much effect on student understanding. This may be due to the fact that the subsequent laboratory exercises in the ERS 102 curriculum were not inquiry-based. Our density laboratory exercise may have been so different from the other exercises that the students were not able to synthesize what they had learned about density in a meaningful way, or forgot much of what they had learned from the exercise because they were not accustomed to inquiry-based methods of thinking.

(4) The laboratory exercise may have been too unconnected from the assessment. The goals for the laboratory exercise were clearly set at the outset, guided by the

overarching research question. But, our results suggest that there was too large a cognitive leap between the laboratory exercise and the assessment questions, particularly the earth science contextual questions. We had hoped that the laboratory exercise would give the students the tools to understand and transfer their understanding of density to other contexts. However, never explicitly teaching convection (assessment questions 1 and 2), exploring plate tectonics (assessment question 2), or demonstrating the effects of salinity on density (assessment question 3) makes the improvement on question 1 rather amazing, and the poor performances on questions 2 and 3 no surprise. The cognitive leap in earth sciences that we hoped students could demonstrate on Part I of the assessment was apparently too much to ask.

Future research on students' understanding of density in earth sciences should include:

(1) Refining of this and other assessments (i.e., the survey used by Yeend et al., 2002). Researchers in physics education have developed the Force Concept Inventory (FCI) (Hestenes et al, 1992) and other conceptual inventories. It is our opinion that the field of earth sciences needs such a concept inventory (including density), and we hope that the assessments used in this study will be a stepping-stone to such an overarching instrument.

(2) Revision of the assessment developed for this study, concurrently with revision of the density laboratory exercise, so that both may be implemented more usefully in earth science courses. A closer pairing of the laboratory exercise and the pre-/post-course assessment may have a greater positive influence on and better ability to judge students' understanding of density.

(3) Use of larger groups of participants, for more reliable statistical conclusions. The two experimental groups in this study were small, and minor fluctuations among earned rubric levels were not necessarily reliable indicators of the overall group's changing level of understanding of density. Furthermore, the intervention and assessments should be used in introductory-level earth science courses at other institutions, to increase the overall number of participants, and to check for reliability of results. Additionally, any continuation of this study should employ control and experimental groups, to more definitively measure any effects of the intervention.

Despite the shortcomings in the study, and the necessity for additional work on the topic, we learned that the concept of density is one that many college students do not understand; many questions were answered correctly by less than one-half of the students, even after instruction. Surely every student who participated in this study had been exposed to density at many points in their education, though they may not have understood it then. We hoped that our inquiry-based activity allowed all students to explore density, mass, and volume in a fun way (the Silly Putty™ was a big hit with students), which would make the subject more interesting and engaging. Furthermore, each group was charged with the task of designing their own laboratory experiment, with only a small amount of guidance from the instructor. Students played an active role in nearly every aspect of the exercise, and so we expected that many students would take ownership of the concepts within the exercise. Although we did not assess possible gain in understanding the scientific method, the laboratory exercise was designed in that format and may have helped to guide students through the density inquiry. Additionally, the exercise brought the scientific method to life for students who may have only read about it.

Clearly, density is a very important concept, pervading much of science. Current educational practices seem not to have developed understanding in the entire K-13+ experience. More attention is likely necessary in grades K-12, so that students at the college level can be more proficient in the language of density. This would allow for college instructors to spend more class time on complex processes in earth and other sciences, and less class time reviewing the basics.

REFERENCES

Balfe, Carol A. 2005. *Unit 1: Density and Buoyancy: Density, a Fundamental Property of Matter: Concepts, Misconcepts, and Difficulties in Teaching and Learning.* 8th Grade Physical Science Curriculum. 1 July 2005 <http://tlc.ousd.k12.ca.us/~acody/densitymisc.html>.

Coladarci, Theodore, et al. *Fundamentals of Statistical Reasoning in Education.* John Wiley & Sons, 2004.

Delany, Robert. *MLA Citation Style.* MLA Handbook for Writers of Research Papers, 6th edition. 2004. Long Island University. 11 Aug. 2005 <http://www.liu.edu/cwis/cwp/library/workshop/citmla.htm>.

Driver, Rosalind, et al. eds. *Children's Ideas in Science*. Philadephia: Open University Press, 1985.

Driver, Rosalind, et al. *Making Sense of Secondary Science: Research into Children's Ideas*. London: RoutledgeFalmer, 2003.

Hannula, Kimberly A. 2003. *Revising Geology Labs to Explicitly Use the Scientific Method.* Journal of Geoscience Education. **51**(2): 194-200. 3 March 2004 <http://www.nagt.org/files/nagt/jge/abstracts/Hannula_v51n2.pdf>.

Heron, P.R.L., M.E.Loverude, P.S.Shaffer, and L.C.McDermott. 2003. *Helping students develop an understanding of Archimedes' principle. (II.) Development of research-based instructional materials.* Am. J. Phys. **71**(11): 1188-1195.

Hestenes, David, Malcom Wells, and Gregg Swackhamer. 1992. *Force Concept Inventory.* The Physics Teacher. 30 (3): 141-158.

Hewson, Mariana G., and Peter W. Hewson. 1983. *Effect of Instruction Using Students' Prior Knowledge and Conceptual Change Strategies on Science Learning*. Journal of Research in Science Teaching. **20**(8): 731-743.

Libarkin, Julie C., Cynthia D. Crockett, and Philip M. Sadler. 2003. *Density on Dry Land*. The Science Teacher. **70**(6): 46-50.

Loverude, Michael E. Assistant Professor of Physics- professional website. California State University, Fullerton. 29 March 2004 <http://chaos.fullerton.edu/~mloverud>.

Loverude, Michael E., Christian H. Kautz, and Paula R. L. Heron. 2003. *Helping students develop an understanding of Archimedes' principle. (I.) Research on student understanding*. Am.J. Phys. **71**(11): 1178-1187.

McDermott, Lillian C., and the Physics Education Group at the University of Washington. *Physics by Inquiry, Volume 1*. New York: John Wiley & Sons, Inc., 1996.

McKinnon, Joe W. 1971. *Earth Science, Density, and the College Freshman.* Journal of Geological Education. **19**(5): 218-220.

Smith, C. et al. 1997. *Teaching for Understanding: A study of students' preinstruction theories of matter and a comparison of two approaches to teaching about matter and density*. Cognition and Instruction. **15**(3): 317-393.

Stavy, Ruth and Dina Tirosh. *How Students (Mis-)Understand Science and Mathematics: Intuitive Rules*. New York: Teachers College Press, 2000.

Stein, Mary, and Dolores Miller. 1998. *Density Explorations*. The Science Teacher. **65**(2): 45-47.

Stepans, Joseph I., Ronald E. Beiswenger, and Steven Dyche. 1986. *Misconceptions Die Hard*. The Science Teacher. **53**(6): 65 to 69.

Wiggins, Grant and Jay McTighe. *Understanding by Design*. Virginia: Association for Supervision and Curriculum Development, 1998.

Yeend, R.E., M.E.Loverude, and B.L. Gonzalez. S. Franklin et al, Eds. *Student understanding of density: a cross-age investigation.* Proceedings of the 2001 Physics Education Research Conference. Rochester, NY, 2002.

APPENDICES

Appendix A

Human subjects proposal:

Measuring Student Understanding and Application of Density

1. Summary of the Proposal:

Topics in the Earth Sciences are commonly characterized by a set of intricately complex processes, interacting with one another in an equally complex manner. Each process, however complicated, is built upon basic governing principles, one of which is density. For example, mantle convection, at its most fundamental level, is a result of density variations within the mantle. Density is the fundamental driving mechanism for innumerable other Earth Science themes, including subduction and accretion of terranes, convection in active zones of permafrost regions, and thermo-haline circulation around the world. A student may not gain a deep understanding of many such processes without being exposed to the fundamental basis for their existence. This study will investigate college students' understandings of density, and determine if a strong conceptual basis leads to fuller understanding of Earth Science processes.

An inquiry-based laboratory exercise focusing on the concept of density will be developed and integrated into the fall 2004-spring 2005 ERS 102 laboratory sequence at the beginning of the semester. This will give the students a fundamental understanding of density from which to draw for the remainder of the semester. Each density experiment within the laboratory exercise will be referenced in subsequent lectures.

The influence of this instructional modification in ERS 102 will be measured using multiple choice and extended response tests at the beginning and end of the semester, as well as performance on the density laboratory experiment. The test questions will be of two types: (1) questions concerning complicated Earth Science processes, governed by density at their most fundamental level; and (2) questions designed to elicit the degree of conceptual understanding of density. Sample questions of both types are attached. The post-course test and the laboratory exercise will be standard components of the course and all students will be required to complete them. However, students may opt to have their performance excluded from the data used in this project. All students will be asked to take the pre-course test voluntarily, and informed consent will be obtained from every participant by their response to the following question on the pre-course test:

"May we use your responses from course assessments and laboratory exercises, without your name or any other identification attached, for education research that may be published?"

2. Personnel:

- Emily L. Klingler, Master of Science in Teaching candidate, University of Maine
- Dr. Stephen Norton, Professor of Earth Sciences, University of Maine
- The other graduate teaching assistant in ERS 102 will be administering the density laboratory exercise and assisting Prof. Norton in grading the course assessments. This person is not yet identified, but when the teaching assistantships are assigned, s/he will be added to this personnel list and will take the IRB tutorial.

3. Participant recruitment:

As many as 240 students (~80/semester for 2 or 3 semesters) in ERS 102 will comprise the population of the study. At the beginning of the course, students will be informed of this study and provided the Informed Consent form (attached).

4. Informed consent:

Each person will be informed of rights, risks, benefits, and the confidentiality level associated with the study. Informed consent will be obtained from every participant by their response to the following question on the pre-course test:

"May we use your responses from course assessments and laboratory exercises, without your name or any other identification attached, for education research that may be published? ___________ (Y/N)"

Signatures for consent will not be obtained.

5. Confidentiality:

Need for pre-/post-course assessment correlation dictates linking the data with human subjects. However, confidentiality will be maintained by the use of codes, with no use of names or University-related identifications. Assessment responses will be kept locked in Emily Klingler's office in the Bryand Global Sciences Center on the University of Maine campus. All electronic files will be stored on the server in the building and will be accessible only to Emily Klingler or Prof. Norton. No identifiers will be used in any publication. Randomly selected pseudonyms will be assigned as necessary. The key to codes and all identifying data will be destroyed within a few months of the completion of the study, leaving only coded data for future use.

6. Risks to Participants:

There is no more risk involved in this study than one encounters in daily life. Each participant's role includes taking a pre-/post-course assessment and participating in the laboratory exercise.

7. Benefits:

The study will shed light on the degree to which an inquiry-based activity in a laboratory setting helps students make connections among concepts and Earth processes. If the laboratory exercise proves useful, it will be permanently adopted in the Earth Sciences and will be available to other teachers. There is little education research in the Earth Sciences on student understanding of density. This study will contribute to building a knowledge base for future reference.

Appendix B

Informed Consent Form

You are invited to participate in a research project being conducted by Emily Klingler, a graduate student in the Master of Science in Teaching program at the University of Maine. The purpose of the research is to investigate understanding of various concepts and to see how effectively those concepts are used in connection with Earth Science topics.

What will you be asked to do?

The data for this study will mainly come from the regular assessments (laboratory exercises, tests, and the final exam) given during the semester. These assessments are used by the professor to determine the students' grades for the course, so you should take them very seriously. You'll also be asked to take one additional assessment at the beginning of the semester that will not influence your grade for the course. Because this first assessment is very important to this study, we ask that you take it just as seriously as you would a regular test. Although your name will be present on all of the assessments, no personal identification (i.e.,, your name or University identification number) will be used in connection with the data.

Risks

Except for your time and the effort in completing one additional assessment (at the beginning of the semester), there are no foreseeable risks to you in participating in this study.

Benefits

The study will help show how well hands-on activities in a laboratory setting help students make connections between concepts and Earth processes. If the laboratory exercise designed for this project proves to be useful, it will be adopted as part of the course and made available to other teachers.

Confidentiality

Although your name will be on all of the assessments, your name will not be on any of the documents developed in this study. A code number will be used to protect your identity. Data will be kept in Emily Klingler's locked office, and only Emily Klingler and Prof. Norton will have access to that data. The key linking your name to the coded data will be destroyed after data analysis is complete. The anonymous coded data will be kept for future use.

Voluntary

You are expected to take all of the assessments in the course very seriously. However, your participation in this study is voluntary. You will have the opportunity on the first assessment to indicate if your responses on the assessments may be used as data in this study. If you choose to let your responses be used in this study, you may change your mind at any time during the semester. Simply contact any of the researchers listed below and ask that your responses be omitted from the study. Choosing to participate, or not to participate, in the study will have no bearing on your grade for the course.

This is the permission question that will appear on the first assessment. Answering "yes" at that time will add your responses to the data for this study.

May we use your responses from course assessments and laboratory exercises, without your name or any other identification attached, for education research that may be published? _______ (Y/N)

Contact information

If you have any questions or concerns about this study, please contact Emily Klingler at:

204 Bryand Global Sciences Center, University of Maine, Orono, ME 04469
Ph: 207-581-1998
Emily_Klingler@umit.maine.edu

You may also reach the faculty advisor for this project, Prof. Stephen Norton, at:
314 Bryand Global Sciences Center, University of Maine, Orono, ME 04469
Ph: 207-581-2156
Norton@maine.edu

If you have any questions concerning your rights as a participant, please contact Gayle Anderson, Assistant to the University of Maine's Protection of Human Subjects Review Board, at:
Ph: 207-581-1498, or
Gayle.Anderson@umit.maine.edu

Appendix C

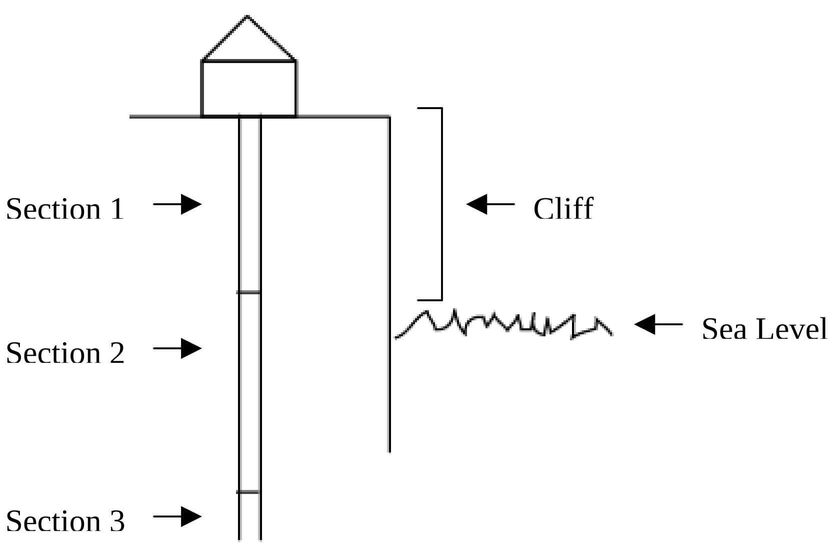
Baseline data (Group 1) questions

1. What process in Earth's mantle is thought to cause plate tectonics? Explain the process fully.

2. When an oceanic plate collides with a continental plate:
(circle which of the following events would occur)

    a. the continental plate is forced downward beneath the oceanic plate, forming a line of volcanoes.
    b. the oceanic plate is forced downward beneath the continental plate, forming a line of volcanoes.
    c. both plates are forced upward, forming a mountain range.
    d. both plates are forced downward, forming a deep trench.

3. A water well was drilled at a popular seaside resort in Maine. As shown in the picture (next page), the resort is on a 50 ft high bedrock cliff, adjacent to the ocean. The driller's record showed three sections as the cliff was drilled. The three sections contained salt water, fresh water, and no water. Predict the order in which these regions would be encountered in the well in the cliff, from highest (encountered first) to lowest (encountered last). Explain why you predicted this pattern.
Sections: salt water, fresh water, no water.



<u>Explain</u> your reasoning:

Appendix D

Fall 2004 (Group 2) pre-course assessment

ERS 102 Assessment

May we use your responses from course assessments and laboratory exercises, without your name or any other identification attached, for education research that may be published? _______ (Y/N)

Part I

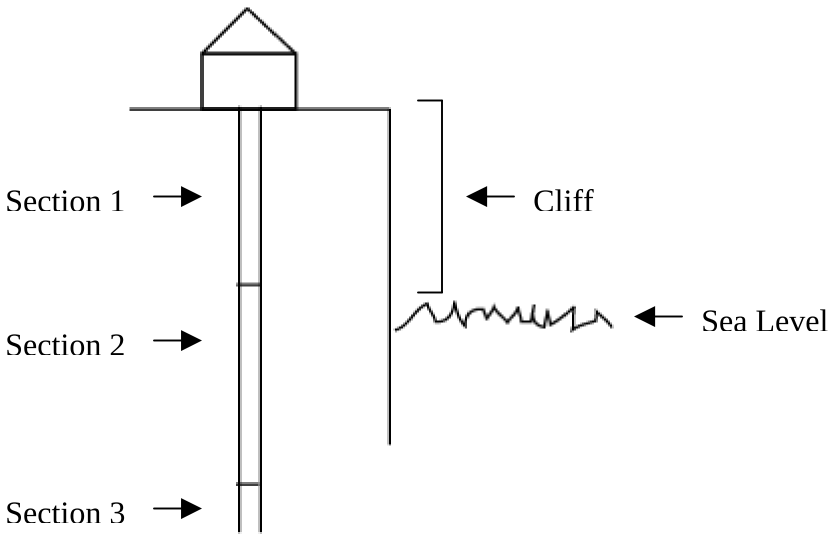
1. What process in Earth's mantle is thought to cause plate tectonics? Explain the process fully.

2. When an oceanic plate collides with a continental plate:
(circle which of the following events would occur)

    a. the continental plate is forced downward beneath the oceanic plate, forming a line of volcanoes.
    b. the oceanic plate is forced downward beneath the continental plate, forming a line of volcanoes.
    c. both plates are forced upward, forming a mountain range.
    d. both plates are forced downward, forming a deep trench.

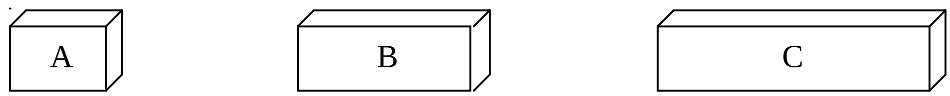
    Explain your reasoning:

3. A water well was drilled at a popular seaside resort in Maine. As shown in the picture (next page), the resort is on a 50 ft high bedrock cliff, adjacent to the ocean. The driller's record showed three sections as the cliff was drilled. The three sections contained salt water, fresh water, and no water. Predict the order in which these regions would be encountered in the well in the cliff, from highest (encountered first) to lowest (encountered last). Explain why you predicted this pattern.
   Sections: salt water, fresh water, no water.



Part II

The following description and pictures apply to questions 4 to 6: A straight, uniform board is cut into three differently sized pieces. Each piece has identical width and thickness, but different lengths.



4. Which piece has the greatest volume?
    a. Piece A
    b. Piece B
    c. Piece C
    d. They are all the same
    e. Impossible to tell without making a measurement.

5. Which piece has the greatest density?
    a. Piece A
    b. Piece B
    c. Piece C
    d. They are all the same
    e. Impossible to tell without making a measurement.

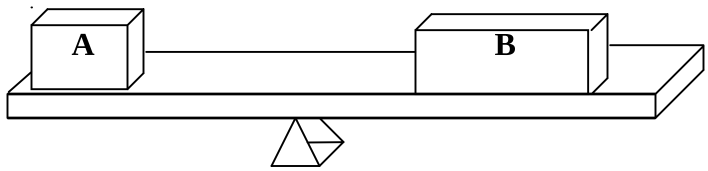
6. Which piece has the greatest mass?
    a. Piece A
    b. Piece B
    c. Piece C
    d. They are all the same
    e. Impossible to tell without making a measurement.

Explain your reasoning for your answers to 4, 5, and 6:

7. A jeweler cut a small chip off a large, uncut diamond. How does the density of the chip compare with the density of the original diamond?
    a. The density of the chip is the same as the density of the original diamond.
    b. The density of the chip is lower than the density of the original diamond.
    c. The density of the chip is higher than the density of the original diamond.
    d. Impossible to tell without taking a measurement.

Explain your reasoning:

The following picture applies to questions 8-9. Objects A and B are uniform rectangular solids, balanced with their centers at equal distances from the pivot point.

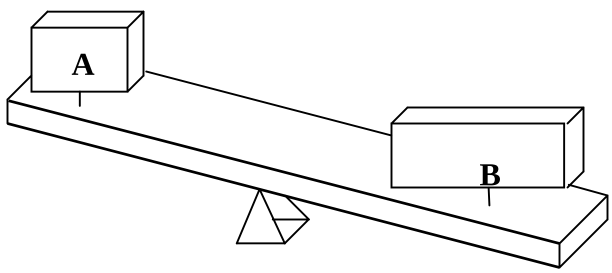


8. Which rectangle has a greater mass for its volume? ______________________________

9. Which rectangle has a greater density?

____

The following picture applies to question 10. Objects A and B are both rectangular solids and are placed on the pans of a balance. The pans are placed at equal distances from the pivot point.



10. Which object has a greater density? ______________________

Explain your reasoning:

Appendix E

Fall 2004 (Group 2) post-course assessment/

Spring 2005 (Group 3) pre-/post-course assessment

Note: There were no changes to the assessment from the post-course version used in the fall of 2004 to the spring 2005. The pre-course assessment had the permission question at the top of the first page, and the post-course assessment did not. Otherwise, the assessments were identical. Only the pre-course version is included here for that reason.

Name:___________________

ERS 102 Post-Course Assessment

Part I

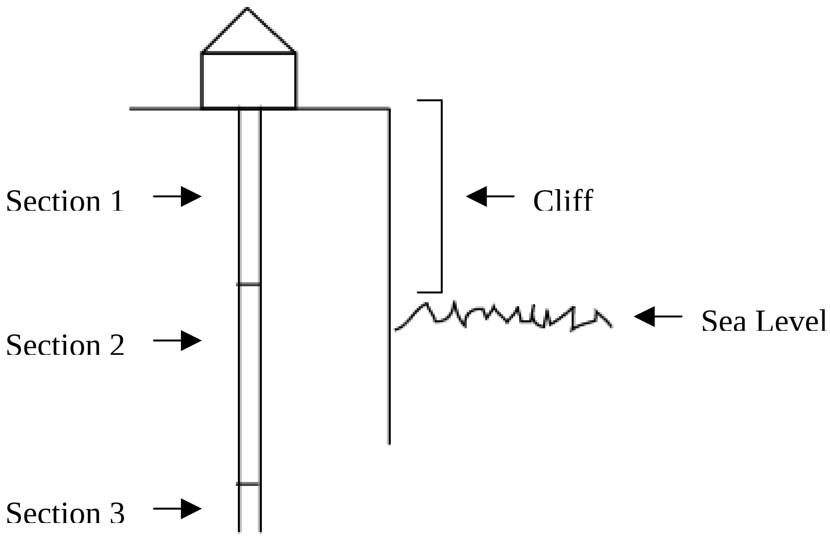
1. What process in Earth's mantle is thought to cause plate tectonics? Explain the process fully.

2. When an oceanic plate collides with a continental plate:
   (circle which of the following events would occur)

A. the continental plate sinks beneath the oceanic plate, forming a line of volcanoes.
B. the oceanic plate sinks beneath the continental plate, forming a line of volcanoes.
C. both plates are forced upward, forming a mountain range.
D. both plates are forced downward, forming a deep trench.

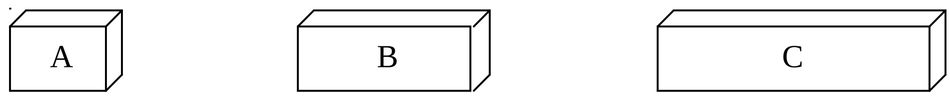
Explain your reasoning:

3. A water well was drilled parallel to a 50 ft high bedrock cliff, adjacent to the ocean, for a popular seaside resort in Maine (as shown in the picture). The driller's record showed three sections as the well was drilled. The three sections contained salt water, fresh water, and no water. Predict the order in which these regions would be encountered in the well, from highest (encountered first) to lowest (encountered last). Explain why you predicted this pattern.
Sections: salt water, fresh water, no water.



Part II

The following description and pictures apply to questions 4 to 6: A straight, uniform board is cut into three differently sized pieces. Each piece has identical width and thickness, but different lengths. A is shortest, C is longest.



4. Which piece has the greatest volume?
    a. Piece A
    b. Piece B
    c. Piece C
    d. They are all the same
    e. Impossible to tell without making additional measurements.

5. Which piece has the greatest density?
    a. Piece A
    b. Piece B
    c. Piece C
    d. They are all the same
    e. Impossible to tell without making additional measurements.

6. Which piece has the greatest mass?
    a. Piece A
    b. Piece B
    c. Piece C
    d. They are all the same
    e. Impossible to tell without making additional measurements.

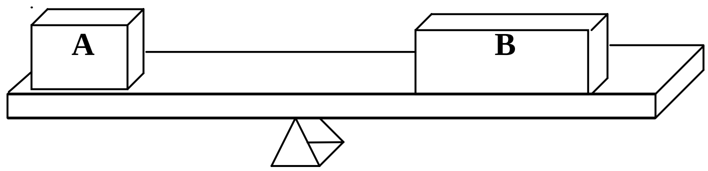
Explain your reasoning for your answers to 4, 5, and 6:

7. A jeweler cut a small chip off a large, uncut diamond. How does the density of the chip compare with the density of the original diamond?

    a. The density of the chip is the same as the density of the original diamond.
    b. The density of the chip is lower than the density of the original diamond.
    c. The density of the chip is higher than the density of the original diamond.
    d. Impossible to tell without taking a measurement.

Explain your reasoning:

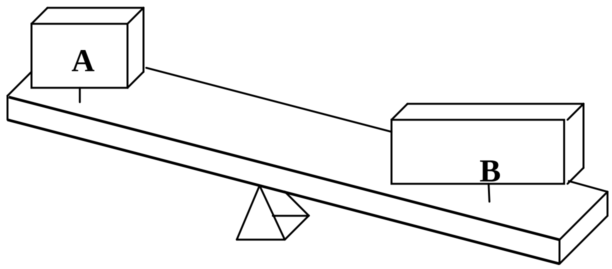
The following picture applies to questions 8-9. Objects A and B are uniform rectangular solids, balanced with their centers at equal distances from the pivot point.



8. Which object has a greater mass for its volume? ______________________

9. Which object has a greater density? ______________________

The following picture applies to question 10. Objects A and B are both rectangular solids and are placed on the pans of a balance. The pans are placed at equal distances from the pivot point.



10. Which object has a greater density? ______________

Explain your reasoning:

Appendix F

Fall 2004 (Group 2) Density laboratory exercise

Laboratory #1: Density

Please find your assigned group for this exercise.
Your group will be given the following equipment:

1. Triple beam balance
2. Graduated cylinder
3. Large beaker
4. Silly Putty, bbs, or a mixture of the two
5. Eyedropper

As you work through the steps in this exercise, please keep a write-up, neat and using complete sentences, which you will hand in at the end of class. Each step below specifies what must be included in the write-up. Feel free to elaborate!

Part 1.

Each group will have 45 minutes for this portion of the lab, after which time each group will present their findings to the class.

A. As a group, discuss the concept of density. In particular, discuss what it is, what it means, and how you might measure it for irregularly shaped objects. In your laboratory write-up, include an interpretation of density that does not include the words “per,” “divided by,” or “over.”

B. Discuss with your group how you would determine the density of water. Try your method and report your findings to the instructor.

C. Create a general procedure for determining the density of an irregularly shaped object. Use several trials of your procedure to “get the bugs out,” then write out your procedural steps in clear definitive language in your laboratory write-up. Also create neat data Tables in your write-up, capable of recording any and all necessary data and/or observations for ten official density measurements.

D. If using Silly Putty or the mixture, break the sample you were given into five pieces. Find the density of each piece twice. Use two different shapes of Putty for each of the five pieces (i.e., use the first piece of Putty for trial #1. Mold it into one shape and take your measurements. Next, use the same piece of Putty for trial #2, but mold it into a different shape. For trial #3, use the second piece of Putty and mold it into one shape, and so on). If you're in the bbs group, simply use different numbers of bbs for your ten density measurements. Record the data in your write-up.

E. When all groups are done, or approximately at the 45-minute mark, each group will present their procedure and experimental results to the rest of the class. (5 min. per group)

F. Summarize the findings for the determination of the density of the Silly Putty, the bbs, and the Silly Putty mixture in the form of three "rules," and include these rules and an explanation of each one in your laboratory write-up.

Intermission—As a group, if necessary (or recommended), revise your procedure and/or data-keeping strategy based on the presentations of other groups. Read ahead through Part 2 of the exercise and create the data Tables necessary for you to complete it using the procedure you developed. Take a quick (3-5 minute) break.

Part 2.

Each group has the remainder of the laboratory time for Part 2 of the exercise.

G. Have your instructor review your procedure and data Tables with you.

H. The instructor will give your group three samples of the same mineral. Record a prediction (to be referenced in your laboratory write-up):

   Will the densities of the mineral samples follow one or more of the "rules" you identified in step F? Explain.

I. Use your procedure from Part 1 to determine the density of each of the three mineral samples. Decide whether the minerals follow any of the rules that you wrote in step F, and explain in your laboratory write-up.

   Are your results consistent with your prediction from step H?

J. Show your results to the instructor. You may not continue until s/he has checked and approved of your group's work.

K. Now you will receive three samples of the same rock from the instructor. Repeat steps G through J with the rock samples.

L. Please answer the following questions in your write-up:

1. Do the mineral and rock samples follow the *same* “rule” from part F?
2. What does this imply about how the composition of minerals compares to the composition of rocks?
3. Referring to your results, what might you conclude about two samples of quartz (or feldspar) from “opposite ends of the earth?” Why?

M. Have a final consultation with your group to be sure that everyone has recorded all necessary data, and that all questions from this exercise have been answered. The write-up is due before you leave.

Appendix G

Spring 2005 (Group 3) Density Laboratory Exercise

Laboratory #1: Density

Please find your assigned group for this exercise.
Your group will be given the following equipment:

1. Triple beam balance
2. Graduated cylinder
3. Large beaker
4. Silly Putty, bbs, or a Silly Putty-bb mixture
5. Eyedropper

As you work through the steps in this exercise, please keep a write-up, neat and using complete sentences, which you will hand in at the end of class. Each step below specifies what must be included in the write-up. Feel free to elaborate!

Part 1.

Each group will have 45 minutes for this portion of the laboratory, after which time each group will present their findings to the class.

A. As a group, discuss the concept of density. In particular, discuss what it is, what it means, and how you might measure it for irregularly shaped objects.

    In your laboratory write-up, articulate your understanding of density, and record a mathematical equation for calculating density.

B. Create a general procedure for determining the density of an irregularly shaped object. Use several trials of your procedure to "get the bugs out," then write out your procedural steps in clear definitive language in your laboratory write-up. Also create neat data Tables in your write-up, capable of recording any and all necessary data and/or observations for ten official density measurements.

C. Run your experiment with your assigned material until you have enough data to defend any pattern you see.

D. When all groups are done, or approximately at the 45-minute mark, each group will present to the rest of the class their procedure and defend any trends found in their data. (5 min. per group)

E. Summarize the findings for the determination of the density of the Silly Putty, the bbs, and the Silly Putty-bbs mixture in the form of three “rules,” and include these rules and an explanation of each one in your laboratory write-up.

Intermission—As a group, if necessary (or recommended), revise your procedure and/or data-keeping strategy based on the presentations of other groups. Read ahead through Part 2 of the exercise and create the data Tables necessary for you to complete it using the procedure you developed. Take a quick (3-5 minute) break.

Part 2.

Each group has the remainder of the laboratory time for Part 2 of the exercise.

F. Have your instructor review your procedure and data Tables with you.

G. The instructor will give your group three samples of the same mineral. Record a prediction in your laboratory write-up:

> Will the densities of the mineral samples follow one or more of the “rules” you identified in step E? Explain.

H. Use your approved procedure to determine the density of each of the three mineral samples. Please answer the following questions in your write-up:

   a. Did your minerals follow any of the rules that you wrote in step E?

   b. Are your results consistent with your prediction from step G?

I. Show your results to the instructor. You may not continue until s/he has checked and approved of your group’s work.

J. Now you will receive three samples of the same rock from the instructor. Repeat steps G through I with the rock samples.

K. Please answer the following questions in your write-up:

   1. Do the mineral and rock samples follow the *same* “rule” from part E?
   2. What does this imply about how the composition of minerals compares to the composition of rocks?
   3. Referring to your results, what might you conclude about two samples of quartz (or feldspar) from “opposite ends of the earth?” Why?

Extra Credit: Determine the density of a piece of pumice, and then answer the following questions in your write-up:

3. How did the pumice behave when placed in water?
4. How did the rocks/minerals behave when placed in water?
5. What can explain the differences in behavior between the pumice and the rocks/minerals?
6. How could you experimentally check your answer to question 8?

L. Have a final consultation with your group to be sure that everyone has recorded all necessary data, and that all questions from this exercise have been answered. The write-up is due before you leave.

Appendix H

Pre-/Post-Course Assessment Rubrics

Rubrics adapted from: Wiggins and McTighe, 1998, p. 76-77.

Table H.1: Question 1

Explanation	Example
5. Sophisticated: convection may or may not be named, with a thorough explanation of convection (temperature and density differences cited); fully supported and justified.	"Convecting and reheating of earth's surface- material gets pushed down because of density and reheats." (the student has drawn convection currents in the mantle and a subduction zone to supplement her explanation.)
4. In-depth: convection may or may not be named, with a sufficient explanation of convection (involving temperature differences); going beyond what is obvious or what was explicitly taught; well supported and justified.	"Convection currents: currents of hotter magma circulate up, cooler magma down. Currents at top of magma create forces which move tectonic plates."
3. Developed/superficial: convection/currents/ named and/or cells drawn with related but superfluous information and/or a relevant portion of the rock cycle explained; mechanisms for convection insufficiently explained.	"Forget the name but the mantle flows in a circular pattern causing the plates to almost float above and that floating leads to movement."
2. Intuitive: an incomplete/incorrect account, with apt and insightful ideas; account has sweeping generalizations or limited support; may cite an incorrect mechanism.	"Earthquakes and volcanoes in the core erupting up to the mantle causes separation and formation of giant mass."
1. Naïve: a superficial account; more of a restatement of the question than an explanation; a fragmentary sketch or glib generalization; more of a borrowed idea than a theory; cites an incorrect mechanism.	"The movement of the plate that constitute the Earth system."
0. No response	

Table H.2: Question 2

Explanation	Examples
5. Sophisticated: correct choice and explanation involving density differences between the two plates.	"Oceanic crust is more dense (denser) than the continental crust."
4. In-depth: correct choice with reference to density/weight/mass differences between the two plates, but insufficiently explained.	"The oceanic plate is made of heavier elements such as Iron and Basalt there the lighter continental crust floats over it, while the other crust is pushed under."
3. Developed/superficial: may not have the correct choice with a limited explanation involving incorrect assumptions or backwards relationships for the two plates; may cite weight/mass/density/hardness; may use a "textbook" diagram with no insightful comments.	"The continental plate is larger and heavier than the oceanic plate. When the two collide, the smaller plate is forced underneath." "The oceanic plate is thinner, less mass and is forced down under the continental plate."
2. Intuitive: may not have the correct choice but has insightful ideas, which are incomplete and/or incorrect.	"Oceanic plates sit lower and therefore pass underneath continental plates."
1. Naïve: may not have the correct choice; a superficial account; more of a restatement of the question than an explanation; a fragmentary sketch or glib generalization.	"The oceanic plate always goes beneath the continental."
0. No explanation offered for choice	

Table H.3: Question 3

Explanation	Examples
5. Sophisticated: correct sequence (none, fresh, salt water) and a correct/complete explanation involving density relationships.	(N,F,S) “The fresh water would be above the salt water because the salt water has a higher density, allowing the fresh water to float above it.”
4. In-depth: correct sequence (none, fresh, salt water) and a good attempt at an explanation, involving density/weight/mass of the units; insufficiently explained.	(N,F,S) “Section 1 would be well drained for being on the cliff side. Section 2 is where freshwater would be draining into the ocean. Section 3 because salt water is heavy and would have the pressure to seep into the landscape.”
3. Developed/superficial: may not have correct sequence; may have backwards density relationships and incorrect assumptions.	(N,S,F) “Salt water (more massive than fresh water) pushing fresh water down.” (N,S,F) “Section 3 could contain fresh water because it is more dense causing it to rise to the top. The salt makes the water less dense…”
2. Intuitive: may not have the correct sequence but has insightful ideas which are incomplete/incorrect.	(N,F,S) “The cliff acts as a barrier between the ocean and fresh water. Not until the salt can go under the cliff that the water will enter the well.”
1. Naïve: may not have the correct sequence; a superficial account; more of a restatement of the question than an explanation; a fragmentary sketch or glib generalization.	(N,S,F) “I say no water first because when I think of the word cliff I think ledge, and when I think ledge, I don’t think there would be water. I say salt water next because it is at sea level and the ‘sea’ is salt water. I then put fresh water because it’s the last choice!”
0. No explanation offered for sequence.	

Table H.4: Question 4

Explanation	Examples
5. Sophisticated: correct choice; correct use of V=LWH.	“Volume=LxWxH thus if they are all equal width and height the longest would have the largest volume.”
4. In-depth: correct choice cited as largest/can hold the most “stuff;” volume may be confused with area.	“Volume is the measure of ‘stuff.’”
3. Developed/superficial: may not have correct choice; incorrect use of V=LWH equation.	“I think if they are identical in width and thickness that volume and density would be the same, but because piece c is longer it may have more mass.”
2. Intuitive: may not have the correct choice but has insightful ideas which are incomplete/incorrect.	
1. Naïve: may not have the correct choice; a superficial account; a fragmentary sketch or glib generalization.	“Depends on what you are putting in it.” “The greater the surface area of an object, the less volume it contains.”
0. No explanation offered for choice.	

Table H.5: Question 5

Explanation	Examples
5. Sophisticated: correct choice; clear explanation of “uniform board” with same density throughout; may use density equation.	“They’re all cut from the same piece of wood so the mass per unit volume will be the same.”
4. In-depth: correct choice; possible mention of “uniform board” with no explanation; cites “same material;” no density explanation included but implied.	“…they are all from the same board.” “The are all the same material.”
3. Developed/superficial: may not have correct choice; limited/incomplete explanation; incorrect use of density equation or attempt at applying density.	“I think that if they are identical in width and thickness that volume and density would be the same…” “Density = a relationship of mass and volume.”
2. Intuitive: may not have the correct choice but has insightful ideas, which are incomplete/incorrect (i.e., density is thickness).	“…Impossible because some species of tree have heartwood with significantly denser heartwood.”
1. Naïve: may not have the correct choice; a superficial account; a fragmentary sketch or glib generalization; cites need for specific measurements.	“Density cannot be figured without weight.” “It looks like more is crammed into a smaller space.”
0. No explanation offered for choice.	

Table H.6: Question 6

Explanation	Examples
5. Sophisticated: correct choice; used dv=m correctly.	"D=m/v; lwh=V. Because c had the largest l, it had the largest V. In order to keep density the same, its m must also be greater."
4. In-depth: correct choice; reference to mass/weight/size but incomplete/absent density explanation.	"…c is the biggest object."' "…because piece c is longer it may have more mass."
3. Developed/superficial: may not have correct choice; limited explanation involving incorrect assumptions or backwards relationships.	"m=lxw. C has greater l than a or b."
2. Intuitive: may not have the correct choice but has insightful ideas, which are incomplete/incorrect.	"…impossible because some species of tree have heartwood with significantly denser heartwood."
1. Naïve: may not have the correct choice; a superficial account; a fragmentary sketch or glib generalization; cites need for specific measurements.	"Without having an exact number, they could all have the same or it could be different." "m = weight."
0. No explanation offered for choice.	

Table H.7: Question 7

Explanation	Examples
5. Sophisticated: correct choice; proper and thorough explanation either: using d=m/v appropriately; involving packing of units; or involving properties of a uniform material.	"The density is weight per volume. The chip has less volume but not weight per volume." "The density doesn't change because density is how tightly packed the molecules that make up the diamond are. So density stays the same even though it is broken into separate pieces." "It came off the original so if the diamond was uniform throughout the density would be equal…"
4. In-depth: correct choice; mention of "same material" in explanation but no use of density equation for support. Not a complete explanation.	"Density is not in size. Density is in material." "A piece of the same item."
3. Developed/superficial: may not have correct choice; possible misuse of density equation or backwards assumptions/ relationships; may confuse area with volume.	"The density (m/v) would be smaller, because of the small mass of the chip."
2. Intuitive: may not have the correct choice but has insightful ideas, which are incomplete/incorrect.	"Diamonds are composed of many parts and can vary in structure. This can cause density changes even when they're all from the same chunk."
1. Naïve: may not have the correct choice; a superficial account; a fragmentary sketch or glib generalization; cites need for specific measurements.	"We can't know for sure how the chip will be without know[ing] the weight and volume of it."
0. No response.	

Table H.8: Question 10

Explanation	Examples
5. Sophisticated: accepTable response (not enough information; either; the same); explains the need for d=m/v for each block.	"B obviously has more mass, but it obviously has more volume as well. If it had equal mass or equal volume, density could be compared."
4. In-depth: correct response; may cite a need to know the blocks' material; no explicit density explanation; not thorough enough.	"Density doesn't mean it has more mass." "Block B has the same density as block A, but is has a greater mass."
3. Developed/superficial: may not have correct response; explanation may attempt using density but incorrect/backwards assumptions (i.e., bigger=denser); may use density equation incorrectly.	"Because B is heavier, density is greater." "More density creates a 'heavier' solid." "D=mv. Either A or B could be more dense. Need more info."
2. Intuitive: may not have the correct response but has insightful ideas, which are incomplete/incorrect.	"A, for the fact that it is smaller, and probably has more 'stuff,' stuffed in it. B just weighs the scale down because it has more mass."
1. Naïve: may not have the correct response; a superficial account; a fragmentary sketch or glib generalization; cites need for specific measurements.	"The scale tips in B's favor, that is my assumption." "[A] It isn't weighing down its side."
0. No response.	

Appendix I

Data tables for statistically insignificant results
Significance level, α=0.05

Table I.1a: Cross-tabulation for post-course question 1, Groups 1 and 2

	Percent within Group 1	Percent within Group 2	Total (n)
0	5%	9%	6
1	6	16	9
2	27	22	25
3	38	28	34
4	20	35	21
5	5	0	3
Total (n)	66	32	98

Table I.1b: Results for post-course question 1, Groups 1 and 2

	Value	df	p value
Pearson Chi-Square	6	5	0.353

Table I.2a: Cross-tabulation for post-course question 2, Groups 1 and 2

	Percent within Group 1	Percent within Group 2	Total (n)
0	12%	6%	10
1	15	16	15
2	15	25	18
3	17	16	16
4	6	6	6
5	35	31	33
Total (n)	66	32	98

Table I.2b: Results for post-course question 2, Groups 1 and 2

	Value	df	p value
Pearson Chi-Square	2	5	0.854

Table I.3a: Cross-tabulation for post-course question 3, Groups 1 and 3

	Percent within Group 1	Percent within Group 3	Total (n)
0	3%	6%	6
1	49	51	65
2	14	9	15
3	0	8	5
4	0	2	1
5	35	25	39
Total (n)	66	65	131

Table I.3b: Results for post-course question 3, Groups 1 and 3

	Value	df	p value
Pearson Chi-Square	9	5	0.129

Table I.4a: Cross-tabulation of pre-/post-course question 2, Group 2

	Post-02						
Pre-02	0	1	2	3	4	5	Total (n)
0	0%	0%	39%	23%	15%	23%	13
1	20	60	20	0	0	0	5
2	0	17	33	0	0	50	6
3	17	17	0	33	0	33	6
4	0	0	0	0	0	100	1
5	0	0	0	0	0	100	1
Total (n)	2	5	8	5	2	10	32

Table I.4b: Results for pre-/post-course question 2, Group 2

	Value	df	p value
Pearson Chi-Square	28	25	0.315

Table I.5a: Cross-tabulation of pre-/post-course question 4, Group 2

	Post-04					
Pre-04	0	1	3	4	5	Total (n)
0	20%	0%	0%	60%	20%	5
3	0	33	33	0	33	3
4	9	9	0	54.5	27	11
5	15	0	8	8	69	13
Total (n)	4	2	2	10	14	32

Table I.5b: Results for pre-/post-course question 4, Group 2

	Value	df	p value
Pearson Chi-Square	20	12	0.068

Table I.6a: Cross-tabulation of pre-/post-course question 10, Group 2

	Post-10						
Pre-10	0	1	2	3	4	5	Total (n)
0	20%	0%	0%	40%	10%	30%	10
1	25	0	25	0	25	25	4
2	0	0	0	50.0	50.0	0	2
3	20	20	0	40	20	0	5
4	0	0	0	14	43	43	7
5	0	25	0	0	0	75	4
Total (n)	4	2	1	8	7	10	32

Table I.6b: Results for pre-/post-course question 10, Group 2

	Value	df	p value
Pearson Chi-Square	28	25	0.315

Table I.7a: Cross-tabulation of pre-/post-course question 2, Group 3

	Post-02						
Pre-02	0	1	2	3	4	5	Total (n)
0	31%	25%	19%	19%	0%	6%	16
1	10	20	35	25	5	5	20
2	4	26	35	13	9	13	23
3	0	0	50	0	0	50	2
4	0	0	100	0	0	0	1
5	0	33	0	33	0	33	3
Total (n)	8	15	20	12	3	7	65

Table I.7b: Results for pre-/post-course question 2, Group 3

	Value	df	p value
Pearson Chi-Square	20	25	0.724

Table I.8a: Cross-tabulation of pre-/post-course question 5, Group 3

	Post-05						
Pre-05	0	1	2	3	4	5	Total (n)
0	29%	14%	0%	14%	43%	0%	7
1	0	19	0	31	38	13	16
2	0	0	0	100	0	0	1
3	8	8	0	31	39	15	13
4	5	9	5	0	73	9	22
5	0	0	0	0	67	33	6
Total (n)	4	7	1	11	34	8	65

Table I.8b: Results for pre-/post-course question 5, Group 3

	Value	df	p value
Pearson Chi-Square	31	25	0.201

Table I.9a: Cross-tabulation of pre-/post-course question 10, Group 3

	Post-10						
Pre-10	0	1	2	3	4	5	Total (n)
0	11%	11%	11%	22%	33%	11%	9
1	6	13	0	38	25	19	16
2	0	0	0	57	29	14	7
3	0	8	0	33	17	42	12
4	0	17	0	25	50	8	12
5	0	0	0	0	11	78	9
Total (n)	2	6	1	19	19	18	65

Table I.9b: Results for pre-/post-course question 10, Group 3

	Value	df	p value
Pearson Chi-Square	33	25	0.141

Table I.10a: Cross-tabulation of post-course questions 5 and 7, Group 2

	Post-07						
Post-05	0	1	2	3	4	5	Total (n)
0	50%	0%	0%	0%	25%	25%	4
1	0	50	0	0	0	50	2
2	0	0	0	0	0	100	1
3	0	0	0	0	100	0	1
4	18	0	0	0	46	36	11
5	0	0	8	8	39	46	13
Total (n)	4	1	1	1	12	13	32

Table I.10b: Results for post-course questions 5 and 7, Group 2

	Value	df	p value
Pearson Chi-Square	29	25	0.261

Table I.11a: Cross-tabulation of pre-course questions 8 and 9, Group 2

	Pre-09		
Pre-08	a*	b	Total (n)
a*	67%	33	15
b	71	29	17
Total (n)	22	10	32

Table I.11b: Results for pre-course questions 8 and 9, Group 2

	Value	df	p value
Pearson Chi-Square	0.06	1	0.811

Table I.12a: Cross-tabulation of post-course questions 8 and 9, Group 2

	Post-09		
Post-08	a*	b	Total (n)
a*	85%	15%	20
b	75	25	12
Total (n)	26	6	32

Table I.12b: Results for post-course questions 8 and 9, Group 2

	Value	df	p value
Pearson Chi-Square	0.5	1	0.483

Table I.13a: Cross-tabulation of post-course questions 4 and 6, Group 2

	Post-06				
Post-04	0	1	4	5	Total (n)
0	25%	0%	50%	25%	4
1	50	0	50	0	2
3	50	0	50	0	2
4	0	10	70	20	10
5	14	0	57	29	14
Total (n)	5	1	19	7	32

Table I.13b: Results for post-course questions 4 and 6, Group 2

	Value	df	p value
Pearson Chi-Square	9	12	0.741

Appendix J

Data tables for multiple-choice portions of the assessment
Significance level, α=0.05

Table J.1a: Cross-tabulation of pre-/post-course m. c. question 2, Group 2

	Post-02				
Pre-02	a	b*	c	d	Total (n)
a	**0%**	100%	0%	0%	1
b*	8	**92**	0	0	25
c	33	33	**33**	0	3
d	0	100	0	**0**	2
No response	0	100	0	0	1
Total (n)	3	28	1	0	32

Table J.1b: Results for pre-/post-course m. c. question 2, Group 2

	Value	df	p value
Pearson Chi-Square	13	8	0.110

Table J.2a: Cross-tabulation of pre-/post-course m. c. question 4, Group 2

	Post-04					
Pre-04	a	b	c*	d	e	Total (n)
a	**0%**	0%	0%	100%	0%	1
b	0	**0**	0	0	0	0
c*	4	0	**92**	0	4	28
d	0	0	50	**50**	0	2
e	100	0	0	0	**0**	1
Total (n)	2	0	27	2	1	32

Table J.2b: Results for pre-/post-course m. c. question 4, Group 2

	Value	df	p value
Pearson Chi-Square	39	9	0.000

Table J.3a: Cross-tabulation of pre-/post-course m. c. question 5, Group 2

	Post-05					
Pre-05	a	b	c	d*	e	Total (n)
a	**0%**	0%	0%	67%	33%	3
b	0	**0**	0	0	0	0
c	0	0	**0**	0	0	0
d*	4	0	4	**92**	0	26
e	33	0	0	33	**33**	3
Total (n)	2	0	1	27	2	32

Table J.3b: Results for pre-/post-course m. c. question 5, Group 2

	Value	df	p value
Pearson Chi-Square	14	6	0.029

Table J.4a: Cross-tabulation of pre-/post-course m. c. question 6, Group 2

	Post-06					
Pre-06	a	b	c*	d	e	Total (n)
a	**0%**	0%	0%	0%	0%	0
b	0	**0**	0	0	0	0
c*	4	0	**92**	4	0	26
d	0	0	100	**0**	0	1
e	0	0	40	40	**20**	5
Total (n)	1	0	27	3	1	32

Table J.4b: Results for pre-/post-course m. c. question 6, Group 2

	Value	df	p value
Pearson Chi-Square	13	6	0.044

Table J.5a: Cross-tabulation of pre-/post-course m. c. question 7, Group 2

	Post-07					
Pre-07	a*	b	c	d	No Response	Total (n)
a*	**89%**	4%	0%	4%	4%	27
b	50	**0**	0	50	0	2
c	100	0	**0**	0	0	1
d	0	0	0	**0**	0	0
No response	100	0	0	0	**0**	2
Total (n)	28	1	0	2	1	32

Table J.5b: Results for pre-/post-course m. c. question 7, Group 2

	Value	df	p value
Pearson Chi-Square	7	9	0.601

Table J.6a: Cross-tabulation of pre-/post-course m. c. question 2, Group 3

	Post-02					
Pre-02	a	b*	c	d	No response	Total (n)
a	**25%**	75%	0%	0%	0%	4
b*	10	**66**	14	7	50	29
c	6	59	**18**	12	50	17
d	14	71	0	**14**	0	7
No response	0	50	38	13	**0**	8
Total (n)	6	41	10	6	2	65

Table J.6b: Results for pre-/post-course m. c. question 2, Group 3

	Value	df	p value
Pearson Chi-Square	9	16	0.912

Table J.7a: Cross-tabulation of pre-/post-course m. c. question 4, Group 3

	Post-04					
Pre-04	a	b	c*	d	e	Total (n)
a	**0%**	0%	0%	0%	0%	0
b	0	**0**	0	0	0	0
c*	3	0	**92**	5	0	61
d	0	0	100	**0**	0	3
e	0	0	0	0	**100**	1
Total (n)	2	0	59	3	1	65

Table J.7b: Results for pre-/post-course m. c. question 4, Group 3

	Value	df	p value
Pearson Chi-Square	65	6	0.000

Table J.8a: Cross-tabulation of pre-/post-course m. c. question 5, Group 3

	Post-05					
Pre-05	a	b	c	d*	e	Total (n)
a	**40%**	0%	40%	20%	0%	5
b	0	**0**	0	0	0	0
c	0	0	**20**	80	0	5
d*	2	0	0	**93**	5	43
e	0	0	0	75	**25**	12
Total (n)	3	0	3	54	5	65

Table J.8b: Results for pre-/post-course m. c. question 5, Group 3

	Value	df	p value
Pearson Chi-Square	42	9	0.000

Table J.9a: Cross-tabulation of pre-/post-course m. c. question 6, Group 3

	Post-06					
Pre-06	a	b	c*	d	e	Total (n)
a	**0%**	0%	0%	0%	0%	0
b	0	**0**	0	0	100	1
c*	0	0	**89**	6	6	53
d	0	0	50	**25**	25	4
e	0	0	67	0	**33**	6
Total (n)	0	0	54	4	7	65

Table J.9b: Results for pre-/post-course m. c. question 6, Group 3

	Value	df	p value
Pearson Chi-Square	17	8	0.03

Table J.10a: Cross-tabulation of pre-/post-course m. c. question 7, Group 3

	Post-07				
Pre-07	a*	b	c	d	Total (n)
a*	**98%**	0%	0%	2%	50
b	50	**25**	25	0	8
c	67	33	**0**	0	3
d	100	0	0	**0**	3
No response	100	0	0	0	1
Total (n)	59	3	2	1	65

Table J.10b: Results for pre-/post-course m. c. question 7, Group 3

	Value	df	p value
Pearson Chi-Square	32	12	0.002

BIOGRAPHY OF THE AUTHOR

Emily L. Klingler was born in Norway, Maine, and graduated from Oxford Hills High School. She attended Boston University with a major in Earth Sciences, and conducted field work in Western Ireland and the Dry Valleys, Antarctica while attending university. She received a Bachelor of Arts in Earth Sciences from Boston University in 2003. Emily is a candidate for the Master of Science in Teaching degree from The University of Maine in May, 2006.