Eye Tracking And Prompts For Improved Learning

An Interactive Qualifying Project Report

Submitted to the Faculty

of the

Worcester Polytechnic Institute

In partial fulfillment of the requirements for the

Degree of Bachelor of Science

by Michael Brigham & Eric Levin

Submitted to Dr. Janice Gobert

Department of Social Sciences and Policy Studies

6/1/2012

Abstract

This study was designed to test the effectiveness of a series of automatic prompts provided to students engaged in reading text and viewing images (versus a control group of students not given automatic prompts). Participants were brought in to learn about plate tectonics by the use of microworlds. The prompts were given by a pedagogical agent, a dinosaur named Rex, who directed students to read the text and view images in a more expert-like fashion. Prompts were given automatically based on data from an eye tracking device and software that was written to that track the students' gaze on the screen and provided feedback to them (Gobert & Toto, 2012).

A pre-test that included 4 open response items and 10 multiple choice questions was administered before the microworlds. The same test questions were given then given as a post test as well. The Rex and no Rex groups were compared to see if Rex's automatic scaffolding prompts had an effect on the participants' comprehension and retention (as measured by post-test minus pre-test scores). Based on t-tests run on all pre-test scores, there was not any bias upon entry into either group. Through further t-tests run on the average gains from pre- to post-test, it was concluded that our findings did not support our hypothesis. This may change through further analyses of the eye tracking data, as at this time not all of the data have been examined.

Acknowledgements

Special thanks to:

Dr. Janice Gobert for advising this project and her help with the statistical analysis,

Zakkai Kauffman-Rogoff for his assistance with the ITS and Rex software,

Ermal Toto for his assistance with both software and statistical analysis,

Juelaila Raziuddin for her assistance with the statistical analysis,

and Michael Wixon for his editing.

Contents

Abstract 2

Acknowledgements 3

Introduction 5

Background 7

Hypothesis 10

Goals 10

METHOD 11

Participants 11

Materials 11

Science Assistments system 11

Rex 11

Plate Tectonics unit 12

Eye tracker 12

Intelligent Teaching System 12

Screen tagging for reading and viewing regions 13

Pre and post test questions 14

Data coding of Open Response Data 15

Procedure 16

Results 18

Data analyses 18

Discussion 22

References 24

Appendix A - Screenshots of the Microworlds 26

Appendix B - Pre/Post Test 30

Appendix C - Coding for Open Response Question 1, 2, 3, & 4 33

Introduction

This is a continuation of a Major Qualifying Project carried out by Zakkai Kauffman-Rogoff, “An Intelligent tutoring system with eye tracking-based scaffolding” (Kaufman-Rogoff, 2011). In the previous project, Zakkai along with the help of Dr. Janice Gobert, Dr. David Brown, and Ermal Toto designed a system that used students’ eye-tracking patterns to determine where the students were looking on a screen. Additionally, the system was designed to be a proof of concept that Rex, the pedagogical agent, could be used to direct students’ reading and viewing based on eye tracking patterns. For this proof-of-concept test, Zakkai Kauffman-Rogoff and Ermal Toto, a software engineer, produced the software code for interpreting students’ reading and viewing locations. This software then directed the students’ attention with text messages, delivered by Rex, according to well-defined regions specified by Janice Gobert, head of the Science Assistments group. The goal of these scaffolding messages is to support students' knowledge acquisition processes so as to better use the affordances of each of the media used to convey the material (i.e., text and animations) as measured by changes in their pre and post test scores. It is important to note that we were not able to analyze students’ eye tracking data because time did not permit this. Analyses of these data, however, would be a better index of the efficacy of Rex on knowledge acquisition.

The literature on progressive model building that describes how readers learn from both text and pictured-based learning was used in determining canonical models for the order of the reading and viewing regions. Specific reading and viewing regions were delineated based on how people best develop models of plate tectonic phenomena (Gobert & Clement, 1999; Gobert, 2005).

Geology was the subject chosen for the participants due to its complexity. Geology is a topic that cannot easily be understood with text based materials only (Gobert, 2000). Plate tectonics provide the perfect example of a subject that requires students to view texts and pictures in a proper order in order to understand the phenomena.

When reviewing open response answers it is important to determine if the subject knows the answer or is using key words they remember incorrectly. In a study it was shown that both experts and novices would both use the same amount of words (Jarodzka, 2010), so the length of the answers do not indicate that the answers are more accurate. The information for tectonics was gathered from a grade school (Padilla, Miaoulis, & Cyr, 2009). The animations were created by Gobert on an earlier project (Gobert & Pallant, 2004) with to go alongside the text. Both the text and animations were the source for the test questions as well. Ermal Toto provided aid in developing one of these animations because it was decide to include an addition one to address oceanic-oceanic convergence.

It was important to create materials that would teach the participants with the optimal amount of efficiency. Studies have shown that the comprehension of a subject is increased with both text and diagrams (Hegarty, 1992). How well the images and text are integrated has been known to be a factor in learning (Scheiter & Van Gog, 2009).

It has been shown that effective learning in this domain can be achieved when students are prompted to attend features in the following order: first the spatial (static) features of the domain, second followed by the causal and dynamic features, and finally the plate tectonic phenomena, e.g. mountain formation, volcanic eruption, and sea floor spreading, driven by causal and dynamic processes inside the earth (Gobert & Clement, 1999; Gobert, 2000; Gobert, 2005).

The previous MQP project (Kaufman-Rogoff, 2011) was able to demonstrate that the prompts could be generated and displayed within the microworlds based on the subject's eye movements. While Kaufman-Rogoff was able to create and test the prompts, it was beyond the scope of the previous project to collect data in order to empirically test the efficacy of the pedagogical agent at influencing students' eye tracking, and in turn, comprehension. The goal of the current project is to test the efficacy of the system by post-test over pre-test gains with participants who are from a similar demographic to those for whom the system was designed, i.e., those who in the future may be helped by an intelligent based tutoring system based on eye tracking. Both Kauffman-Rogoff and Toto provided occasional technical assistance in making adjustments to the software for this new project.

The team extended the existing system by developing code for the remaining three microworlds for plate tectonics, and tested out the efficacy of the system with elementary school students. In doing so, this project will provide important data about the efficacy of such a system, a stepping stone for future versions of intelligent based tutors based on eye tracking.

Background

Eye tracking has been used in the past to make inferences about a subject's mode of thinking and how it may be externally influenced. In a study by Barkowsky (2010) attempts were made to identify the differences between the use mental models vs. visual mental images externally influence which mode a subject used.

Graesser, Lu, Olde, Cooper-Pye & Whitten (2005) also used eye tracking data to identify the cognitive processes that occur before, during, and after a subject asked questions about the simulated breakdown of a mechanical model.

Subjects have been observed employing mental imagery to solve causal reasoning problems in a correspondingly systematic fashion. Yoon & Narayanan (2004) suggested that information displays that respond to a user's visual attention trajectory, a kind of Attentive User Interface, were likely to benefit 42% of their study's participants. Further studies by Yoon & Narayanan (2006) suggest that while animation of a procedure does not improve accuracy, animation coupled with progressively revealing objects of interest on the display does improve accuracy and other measures of performance.

The invention of the pedagogical agent Rex was created in order to help students when they are learning during inquiry. This project extends Rex' functionality to supporting students while they are learning from text and pictures, such as is typical when one is reading science materials over the internet. Furthermore, since textbooks are being phased out and many states are beginning to adopt a "one laptop per child" initiative, a system that automatically tracks and scaffolds students' attention based on eye tracking is possible and may benefit their knowledge acquisition processes during on-line reading and viewing of science material.

The eye tracker is the ideal ways to determine what a subject is looking at. It can show what features a subject is focusing on. It can give data on what attracts a subject's attention and if they are paying attention to relevant features (Rex in this case).

It is important to note that eye tracking data can show what the subject is reading but not what they are retaining. Sometimes the human gaze does not directly lead to data (Hyona, 2010). Retention in this study was determined by the participants' change in answers between the pre and post tests. There have been many studies involving eye trackers to collect data that was otherwise unavailable. It can shed light on certain scenarios such as a student reading about a scenario that they cannot directly visualize. With eye tracking, they will look at a visual representation and fixate on the part the text is concerned about (Graesser et al., 2005). This may show what aspects the subject has some understanding about from the material they are reading and thus can also show if they are paying attention to what they are reading/viewing.

Children are the target demographic for this research project because they are the ones who most need help when reading. Students who have the aid of a pedagogical agent may get a head start on learning how to view text and picture based information in a manner that best supports their mental model construction (Gobert, 2005). The hope is that this project will provide empirical data that the prompts provided by Rex will help support students who might not engage in optimal knowledge acquisition processes, for example, those who might otherwise skim or not attend fully to the text, or those who do not use the affordances of each of the representations (text and simulation) to their full advantage.

Rex is an essential part of the project, and the use of a prompts have been important in research in the past. An important and unresolved question that has risen from contradictory findings concerning the effectiveness of pedagogical agents (see e.g. Dehn & Van Mulken, 2000) is whether the agent will draw attention and cognitive resources away from other important information sources on the screen, or whether it will help learners process the information from other sources more effectively. King & Ohya (1996) describe the importance of using a prompt that seems semi intelligent. This suggests the use of an agent like Rex will help more than just a text box in the corner.

Since Rex doesn't just pop out in the front of the screen, it could be that having Rex give oral, as opposed to textual prompts, may improve the students' learning. A study by Van Gog & Scheiter (2010) showed that pictures presented with spoken rather than written text leads to better learning outcomes but a study by Sproull, Subramani, Kiesler, Walker & Waters (1996) showed that a voice makes the prompt unlikeable. The prompt is supposed to give the learner something that will help him/her, without distracting them from the information on the page.

Hypothesis

Our hypothesis is that the agent, Rex, would be able to effectively scaffold learners to read and view in a more expert like fashion, as evidenced by their eye-tracking patterns. Secondly, we hypothesize that due to better knowledge acquisition patterns, comprehension would be better for the Rex condition, as measured by post-test over pre-test gains.

Goals

The prior project (MQP) implemented a system that could be used to identify if a student was not reading/viewing the appropriate sections when compared to a canonical model, and then to provide prompts to learners in order to guide their learning (Kaufman-Rogoff, 2011). The early pilot data from the MQP project demonstrated that Rex had a quick response time that would be activated as soon as the reader veered from the preprogrammed partial order path.

(1) We hope that the use of Rex will help improve students' comprehension of the material, as evidenced by greater post- versus pre-test scores (Rex vs. no Rex). We examined our data both as total pre-post measures as well as disaggregated for spatial versus causal/dynamic understanding. In the current experiment, the participants were unaware that Rex would be used as a tool to help them while they are reading. Since we were testing Rex vs. non Rex conditions, it is appropriate to hide the fact that Rex was part of the test.

(2) We hoped that Rex will be able to attract the attention of the reader without having to point it out to the test subject. Eye tracking patterns for each condition will be analyzed at a later date. These data will help address whether and how the students responded to Rex's scaffolds, even if there are no group differences yielded in the comprehension measures.

METHOD

Participants

This study consisted of 30 volunteer middle school students from central Massachusetts who had no prior classroom exposure to plate tectonics. Each group of participants (5-8 students per session) was put in a drawing for a $15 iTunes gift card as compensation for participation in the study. One student's data had to be eliminated from the data analyses as the eye tracker could not follow their gaze despite multiple calibrations.

Materials

Science Assistments system

WPI's Science ASSISTments project deals with the development of a set of virtual microworlds that allow students to hone science content knowledge and inquiry skills. The ASSISTments project presupposes that engaging students in scientific inquiry processes via microworlds will positively affect students' scientific skills and, in turn, their scientific content knowledge. The science ASSISTments system at assistments.org was used to track participants responses to both pre and post test questions as well as act as the repository for the microworlds.

Rex

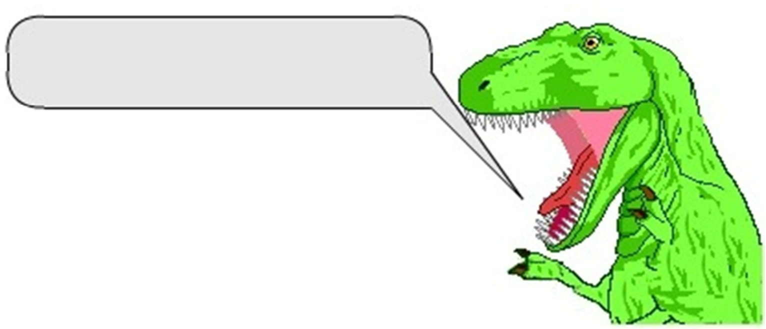


Figure 1 - Rex, the pedagogical agent

The pedagogical agent Rex displays text messages generated by the Intelligent Tutoring System (ITS) eye tracking program from the Kauffman-Rogoff MQP (2011). Rex and his associated speech bubble are always visible to the participants during the Rex condition even when there is no scaffolding message to display. Neither Rex nor his speech bubble were visible during the control condition.

Plate Tectonics unit

Four microworlds were used with this project consisting of the static and causal/dynamic features of plate tectonic phenomena. The microworlds were the layers of the Earth, continental-continental convergence, oceanic-continental convergence, and oceanic-oceanic convergence.

Eye tracker

The type of eye tracker used in this experiment was the Mirametrix S1. This is the same eye tracker device from Kauffman-Rogoff MQP (2011). The eye tracker uses a point-of-gaze from a user's pupils to place where their eyes are on a screen. Calibrations need to be preformed for every subject that uses the eye tracker but can be very accurate when telling where someone is looking on a computer screen. As with most eye trackers, participants with prescription eye glasses tend to generate less accurate readings as many eye trackers cannot distinguish readily between the reflections of the eyes, which they track, and the reflections from the prescription lens.

Intelligent Teaching System

The program used to track participants' eye movements and map them to the corresponding regions on the screen was created by the ITS's region defining software (Kauffman-Rogoff, 2011). Each rectangular screen region contains an area of interest, either text or an image. The reading/viewing areas and their order that a student must attend before the system considers it read/viewed is based on the area defined as actual content rather than the surrounding white space. This is then scaled based on whether the region is designated as image or text. As Rex was originally designed to work with only one microworld per activation, there was a delay between each microworld during testing. The test proctor needed to manually stop the ITS, change which region definition file it was using, load the next microworld and restart the ITS. Zakkai Kauffman-Rogoff was brought in to help the team make adjustments to Rex to accept multiple microworlds without the manual transition, but it was determined to be too great an undertaking in the time allotted. Due to complexity involved it was simpler to have each separate slide run its own instance of the program. This was determined to be a small cost for what would have been a major loss of data had Rex been improperly working during the study. The program is also unable to determine if or when the animated simulations within the microworlds are run.

Screen tagging for reading and viewing regions

The eye tracker is very sensitive to the rapid eye movement of people. The placement of the objects on the screen needed to be arranged such that the eye tracker could easily associate the placement of text and pictures with the participant’s gaze location. These regions inform the program which paragraph or picture that the reader is looking at and determines if all the prerequisite regions have been examined. If not all the prerequisites have been met, such as reading the second paragraph without reading the first, then a scaffolding message is displayed by Rex. A typical message for Rex would be "Please remember to view the first paragraph <and any additional missed items> thoroughly." Each microworld required a unique set of regions boxes around the text and pictures. These regions are what the eye tracker uses to assess which area the reader is looking at and, as such, cannot overlap. See Figure 2 below for an example.

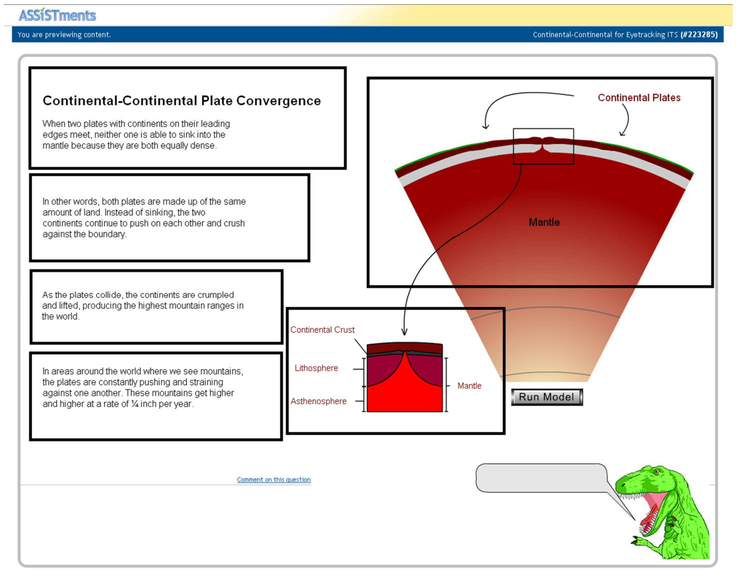


Figure 2 - Sample region definitions

Much of the early work on the microworlds by was determining the proper placement and sizes of images and text to minimize errors in eye tracking. As the eye tracker is not one hundred percent accurate at telling where the viewer is looking, errors in gaze locations could vary up to 2 lines of text (approximately half an inch). Spacing between objects became a critical concern. Having the content portions of the regions too close together caused the eye tracker to confuse which part of the page that reader was actually looking at and resulted in unneeded messages or caused a necessary message to not be displayed. By maximizing the unused space between areas of content we decreased the likelihood that inaccuracies in the eye tracking would lead to false data regarding the participants' gaze location.

Pre and post test questions

The pre and post test consisted of the same 10 multiple choice questions and 4 open response questions. The tests were composed of both static and dynamic questions. Static questions tested the participants' retention of the text while dynamic question were designed to test how much subjects had understood from the corresponding visual aids. A complete copy of the test can be found in Appendix B. The multiple choice was automatically graded by Science ASSISTments.

Data coding of Open Response Data

In order to grade the open response portions of the pre- and post-tests, project advisor Professor Janice Gobert developed a coding key for each question. In brief, the data coding consisted of scoring the extent to which students' open response answers "matched" either by exact recall or paraphrase the text they read for each of the four sections of the text. Similar scoring techniques are used in text comprehension research (Gobert & Clement, 1999). An example of this coding scheme is shown below for oceanic-oceanic plate convergence. The full set for the four open response questions is given in Appendix C.

Table 1 - Open Response Coding for Oceanic-oceanic Convergence

Static/Spatial components S4 (o-o convergence)	**Score Max**
Oceanic Plates (2)	2
Located on floor of ocean	1
Made of basalt	1
Plates are different densities	1
Due to different amounts of basalt	1
Total spatial S4	

Causal/Dynamic components S4 (o-o convergence)	**Score Max**
Plates converge	1
Denser plate slides over the less dense plate	1
One plate sliding over another plate = subduction	1
Subducted plate travels into asthenosphere	1
Subducted plate melts	1
Subducted plate is absorbed into mantle	1
As sinking plate sinks, creates oceanic trench	1
Oceanic trench is in the deepest part of ocean	1
Formation of trenches causes earthquakes	1
Formation of trenches causes volcanoes	1

Volcano is formed from molten material which collects at subduction area	1
Convergence collects molten materials	1
Molten materials contribute to volcanic arcs	1
Volcanic arcs form on top of oceanic plates	1
Volcanic materials collects, forms island	1
Island is located above ocean's surface	1
Total causal & dynamic S4	

Two of the group members then each graded twenty of the students' scores; the 2 coders both scored ten students' data. This overlap was then used to determine inter-rater reliability between the two groups; the inter-rate reliability measures were all between .7 and .9 for the four open response items.

Procedure

Participants were brought to a room where they created an Assistments.org account. These accounts were used for them to take the pretest to assess each participant's prior knowledge of the domain; this was done in group testing situation. Each participant was randomly assigned to either the Rex or control (no Rex) condition and escorted to the eye tracker workstation located in another room. After getting comfortable, the eye tracker was adjusted to account for the height and distance of the subject from the monitor. Participants were reminded to limit the movement of their head as much as possible during the calibration and data collection session to improve accuracy of the eye tracker.

The eye tracker was the calibrated to the individual participant using the software supplied by the manufacturer, Mirametrix. Each participant had an accuracy score that would show how accurate the eye tracker was at linking where the user was looking on the screen to what the program thought the viewer was focusing on. The lower the calibration score meant the more accurate of a reading. Scores below 80 are considered good and below 40 is considered excellent.

Each participant was then recorded for approximately 10 seconds reading the webpage www.thisafterthat.com which was chosen for its large text and spacing to verify calibration. If the calibration was sufficiently inaccurate or the eye tracker was not following the participant's eyes, the calibration process was repeated up to 3 times. The final calibration numbers were recorded for each participant and a note was included if the participant wore glasses.

Due to programming limitations, participants were asked to close their eyes while each new screen was readied to avoid exposure to each microworld's content before Rex was activated. The same procedure was used in the non-Rex condition. The appropriate microworld (Rex/no Rex) was then shown on the screen through the assistments.org website in full screen mode (entered by pressing F11 in Firefox).

In the Rex condition, the ITS program will cover half the page until the space bar is pressed at which time it reveals the current microworld, begins using the data streaming from the eye tracker to map the students' eye movement to the viewing regions and generating the scaffolding messages displayed by Rex. The program stops when the spacebar is pressed again. Students were asked to open their eyes and press the spacebar to begin each microworld and again when they were finished with the page. This spacebar press had no effect during the no Rex condition, but was part of the instructions for consistency between groups. This process was repeated for each of the four microworlds presented following the order of: Layers of the Earth, Continental-Continental Convergence, Oceanic-Continental Convergence, Oceanic-Oceanic

Convergence. The recordings were stopped after the spacebar was pressed to end the final screen.

Results

Data analyses

Analyses of Pre-test data to determine if groups were different from each other BEFORE the intervention.

With the help of Juelaila Raziuddin, Ermal Toto, and Professor Janice Gobert and using SPSS Statistics 17.0 software, several tests were run on the data. In each case, a Levene's test for equality of variances was done first to be sure the distribution of each group's data was not significantly different from each other (if they were, we would be violating the assumptions of the tests whereby we compare the group means to determine differences between the groups). First, a t-test was run on the multiple choice data from the pre-tests in order to ensure that there was not a significant difference between the Rex and no-Rex conditions before the students were assigned to either of the 2 conditions. The results showed that there was not a significant difference between groups on the multiple choice total at pre-test.

Table 2 - Levene's Test for Equality of Variances for Pre-test Multiple Choice Items

F	Sig.
0.088	0.769

These results (see table 2) indicate that the variance of the two groups was not significantly different from each other (p= 0.769), thus, we are not violating the assumptions of the test used to compare the means of the two groups. When comparing the means of the two groups, we found no significant difference between the means (p=0.545); thus, the two groups did not differ from each other at pre-test (see table 3).

Table 3 - T-Test for Equality of Means for Pre-test Multiple Choice Items

Sig. (2-tailed)	Mean Difference	Std. Error Difference
0.545	0.0415	0.0679

T-tests were then run on the results from each of the open response questions on the pre-test, again to insure that there were no significant differences between groups before they were assigned to the Rex/No Rex conditions. These results (see table 4) indicate that the variance of the two groups was not significantly different from each other (p= 0.886), thus, we are not violating the assumptions of the test used to compare the means of the two groups.

Table 4 - Levene's Test for Equality of Variances for Open Response (totals) for Pre-test Items 1-4

Open Response Question at Pre-test (Totals)	F	Sig.
1	0.021	0.886
2	2.832	0.102
3	0.090	0.766
4	0.122	0.729

When comparing the means of the two groups on each of the open response pre-test items (1-4), we found no significant differences between the means of the two groups (p=0.979, 0.522, 0.984, and 0.884); thus, the two groups did not differ from each other at pre-test on any of the four open response pre-test total measures (see table 4).

Table 5 - T-Test for Equality of Means for Open Response (totals) for Pre-test Items 1-4

Question	Sig. (2-tailed)	Mean Difference	Std. Error Difference
1	0.979	0.0637	2.3849
2	0.522	-0.2745	0.4243
3	0.984	0.0131	0.6547
4	0.884	-0.0964	0.6568

Analyses of pre-post gain data to determine if groups were different from each other AFTER the intervention.

T-tests were next run on the average gains from pre-test to post-test for both the multiple choice and open response questions in order to see in there was any significant difference in the gain between groups after the intervention with Rex. First a Levene's test was performed to ascertain whether the variance of the two groups differed. One difference was found for the two groups for Item S3 (see Table 6 below). We proceeded with t-tests of means, as the other three items results yielded no differences on the variances of the groups.

Table 6 - Levene's Test for Equality of Variances for Open Response (totals) for Post-test Items 1-4

Question	F	Sig.
Multiple Choice total	0.873	0.357
Open Response 1	1.127	0.296
Open Response 2	1.820	0.187
Open Response 3	4.668	0.038*
Open Response 4	0.213	0.647

* - statistically significant at the $p < .05$ level of alpha.

We ran a t-test on the two groups, and a significant finding was yielded for open response item 2, favoring the no Rex condition; a borderline significant result was yielded for open response item 1 as well, favoring the no Rex condition. See table 7.

Table 7 - T-Test for Equality of Means for Open Response (totals) for Post-test Items 1-4

Question	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Multiple Choice	0.262	-0.0712	0.0624
1	0.060	2.8807	1.4778
2	0.020*	1.6291	0.6636
3	0.076	1.0098	0.5506
4	0.292	0.9265	0.8605

* - statistically significant at the $p < .05$ level of alpha.

Analyses of pre-post gain data on open response items, broken down by spatial aspects of understanding and causal and dynamic aspects of understanding.

Levene's test for equality of variances was performed, and yielded one difference in the variances on these measures (see 3 C & D, table 8). We proceeded with analyses of groups' means while keeping this caveat in mind.

Table 8 - Levene's Test for Equality of Variances for Spatial, Causal and Dynamic Aspects

Question	F	Sig.
1 Spatial	1.713	0.200
1 Causal and Dynamic	2.086	0.158
2 Spat.	2.222	0.146
2 C & D	2.054	0.161
3 Spat.	0.196	0.661
3 C & D	10.838	0.002 *
4 Spat.	0.000	0.985
4 C & D	3.617	0.066

Analyses were then conducted in order to ascertain whether there were group differences on these measures AFTER the intervention.

A significance difference was yielded on open response 2 causal and dynamic (p=.017) in favor of the no Rex condition. Borderline significance was also found for 2 dependent variables, favoring the No Rex condition, namely, for Item 1 (spatial; p= .056) and item 3 (causal and dynamic; p=.054). See table 9.

Table 9 - T-Test for Equality of Means for Spatial, Causal and Dynamic Aspects

Question	Sig. (2-tailed)	Mean Difference	Std. Error Difference
1 Spatial	0.056	2.3137	1.1684
1 Causal and Dynamic	0.081	0.9592	0.5335
2 Spat.	0.417	0.1111	0.1353
2 C & D	0.017*	1.4624	0.5839
3 Spat.	0.416	0.2160	0.2620
3 C & D	0.054	0.7941	0.3972
4 Spat.	0.551	0.3252	0.5400
4 C & D	0.127	0.6013	0.3842

* statistically significant at the $p < .05$ level of alpha.

Overall, our hypothesis was not supported by our findings in that students in the Rex condition did not, in general, show greater learning gains. Some gains were statistically significant favoring the No Rex condition however; in the next section we address these findings.

Discussion

Though our results did not support our hypothesis, we have several ideas as to why this may be. First, there may have been participants in the control (no Rex) group who naturally read in the expert-like fashion that Rex scaffolded, thus minimizing differences between this group and those who had the Rex condition. Secondly, students in the experimental (Rex) condition may *not* have attended to or followed Rex's instructions (analyses of eye tracking data will confirm or disconfirm this). Thirdly, the students appeared to lack motivation, and many seemed to try to rush through the material without giving much effort, which could skew the results. Another possible situation could be that those in the Rex condition may have been distracted by Rex by becoming too interested in his novelty and in turn becoming disengaged from the task. Finally, not knowing that Rex would be giving them instructions may have lead to some of the participants not paying attention to him. In the future, it may be best to inform the subjects of this before collecting data.

Recordings of each participant’s eye movements were taken but we were not able to analyze them in time for the completion of this project. For future work, examination of these video traces and xml data from students’ eye tracking could address whether the factors listed above are tenable hypotheses about our results.

For future work on this topic, it would be beneficial to get baseline data on the participants' knowledge acquisition strategies, and then divide students into categories to better test the efficacy of Rex to direct students, and subsequently test whether these subjects do better in terms of comprehension. Alterations to the nature of Rex's scaffolding to increase its visibility may help to ensure that participants are aware of and attending to his instructions. Additionally, giving better incentives for good work, limiting the open response questions, and providing a more engaging atmosphere may help to better motivate the participants, thereby providing a better empirical test of the efficacy of Rex. Lastly, performing repeated measures MANOVAs on the pre- and post-test is a better way to analyze these data, but complex statistical analyses such as these are beyond the scope of the project.

References

Barkowsky T. (2010). Diagrams in the Mind: Visual or Spatial? *Diagrams 2010: 1*

Dehn & Van Mulken, (2000). The impact of animated interface agents: a review of empirical research *Int. J. Human-Computer Studies* 52, 1-22

Gobert, J. & Clement, J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching, 36(1), 39-53.*

Gobert, J. (2005). The effects of different learning tasks on conceptual understanding in science: teasing out representational modality of diagramming versus explaining. *Journal of Geoscience Education, 53(4), 444-455.*

Gobert, J. (2000). A typology of models for plate tectonics: Inferential power and barriers to understanding. *International Journal of Science Education, 22(9), 937-977.*

Gobert, J.D., & Pallant, A., (2004). Fostering students' epistemologies of models via authentic model-based tasks. *Journal of Science Education and Technology*. *Vol 13(1),* 7-22.

Gobert, J. & Toto, E. (2012). *An Intelligent Tutoring System with Eyetracking-based Scaffolding. Docket 102.2037-000.*

Graesser A.C., Lu S., Olde B.A., Cooper-Pye E., Whitten S (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory & Cognition 2005, 33(7), 1235-1247*

Hegarty, M. (1992). The mechanics of comprehension and comprehension of mechanics. In K. Rayner (Ed.) *Eye movements and visual cognition: Scene perception and reading*. New York: Springer Verlag.

Hyona J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction 20, 172-176*

Jarodzka H., Scheiter K., Gerjets P & Van Gog T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction 20, 146-154*

Kauffman-Rogoff (2011). An Intelligent tutoring system with eye tracking-based scaffolding. *MQP Worcester Polytechnic Institute.*

King, W. J. & Ohya, J. (1996). The representation of agents: Anthropomorphism, agency, and intelligence. In R. Bilger, S. Guest & M. J. Tauber (Hrsg.), *Human factors in computing systems: Chi'96 electronic conference proceedings.* (Online verfügbar, Stand Ende Juni 2001)

Padilla M.J., Miaoulis, I. & Cyr M. (2009). *Science Explorer: Earth science*. Upper Saddle River, NJ: Prentice Hall.

Scheiter K. & Van Gog T. (2009). Using Eye Tracking in Applied Research to Study and Stimulate the Processing of Information from Multi-representational Sources. *Appl. Cognit. Psychol. 23, 1209–1214*

Sproull, L., Subramani, M., Kiesler, S., Walker, J. H. & Waters, K. (1996). When the interface is a face. *Human-Computer Interaction, 11, 97-124.*

Van Gog T. & Scheiter K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction 20, 95-99*

Yoon, D. & Narayanan, N. H. (2004). Mental imagery in problem solving: An eye tracking study. *Proceedings of the Third ACM Symposium on Eye Tracking Research & Applications*, Association for Computing Machinery (ACM) Press, pp. 77-83.

Yoon, D. & Narayanan, N. H., Lee, S. & Kwon, O. (2006). Exploring the effect of animation and progressive revealing on diagrammatic problem solving, *Diagrammatic Representation*

& Inference: Proceedings of the Fourth International Conference on the Theory and Application of Diagrams, Lecture Notes in Artificial Intelligence, Vol. LNAI 4045, Springer, pp. 226-240.

Appendix A - Screenshots of the Microworlds

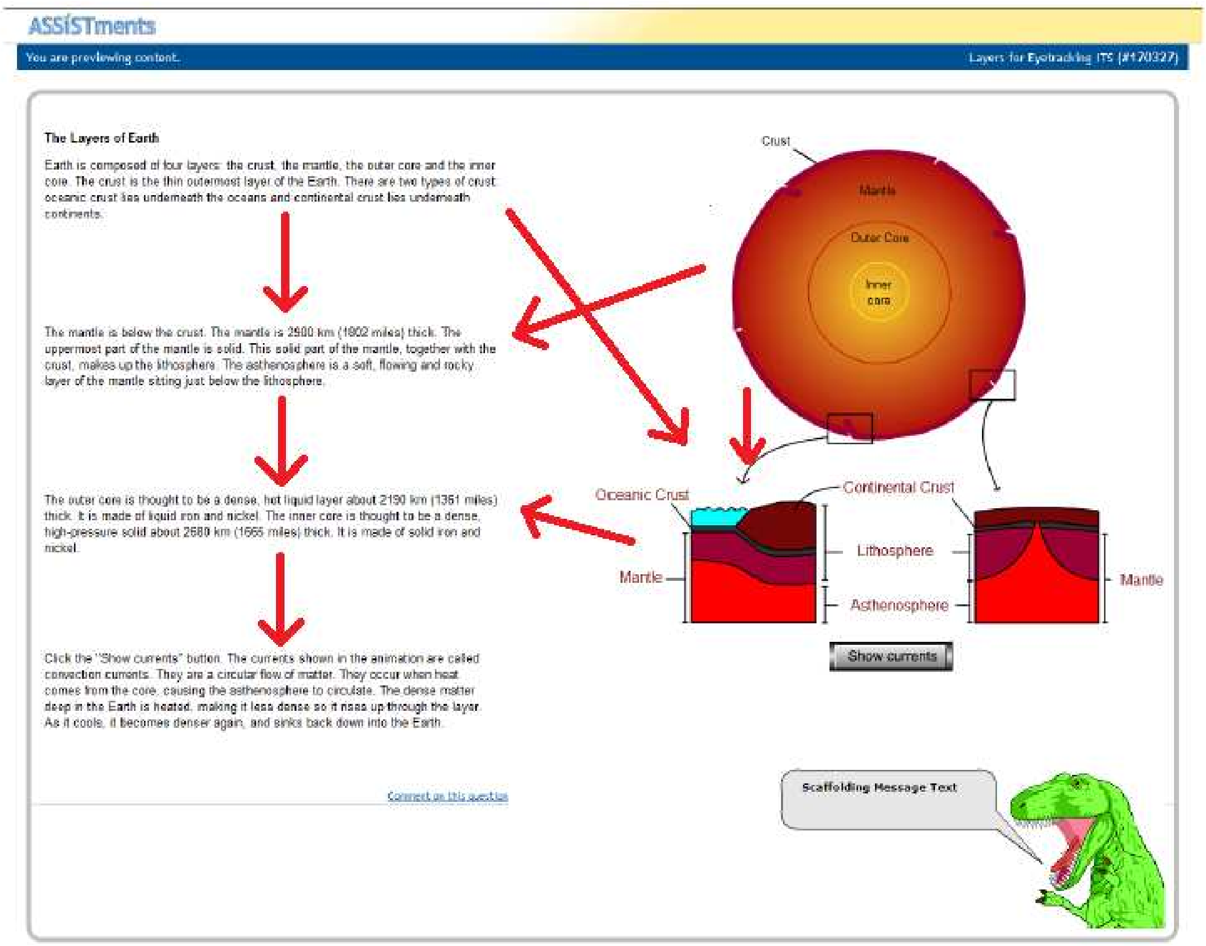


Figure 3 - Image illustrates the partial order list path which would most efficiently aid a reader in comprehension of the material, based on a progressive model-building approach.

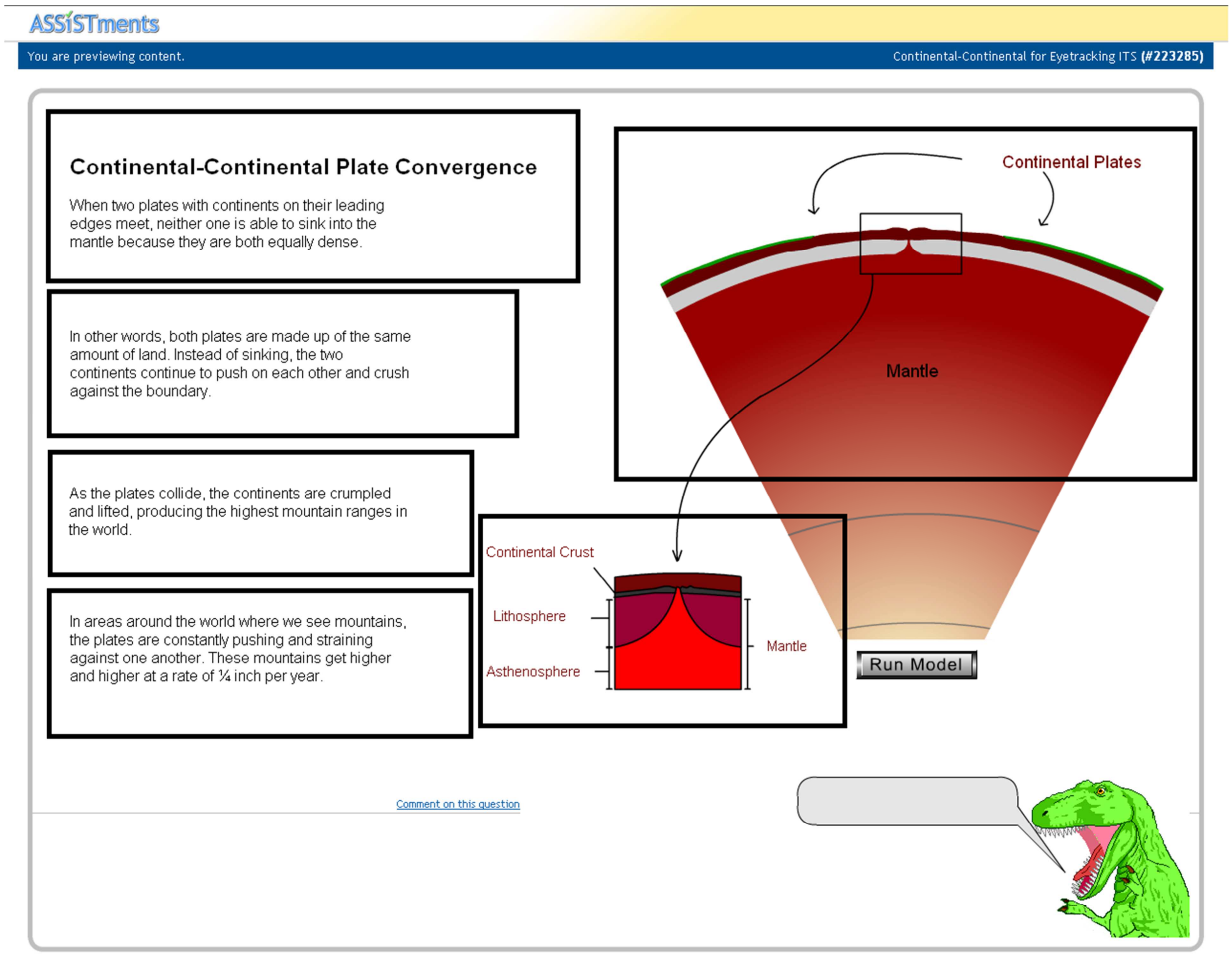


Figure 4 - Image of the continental-continental plate convergence microworld with defined regions

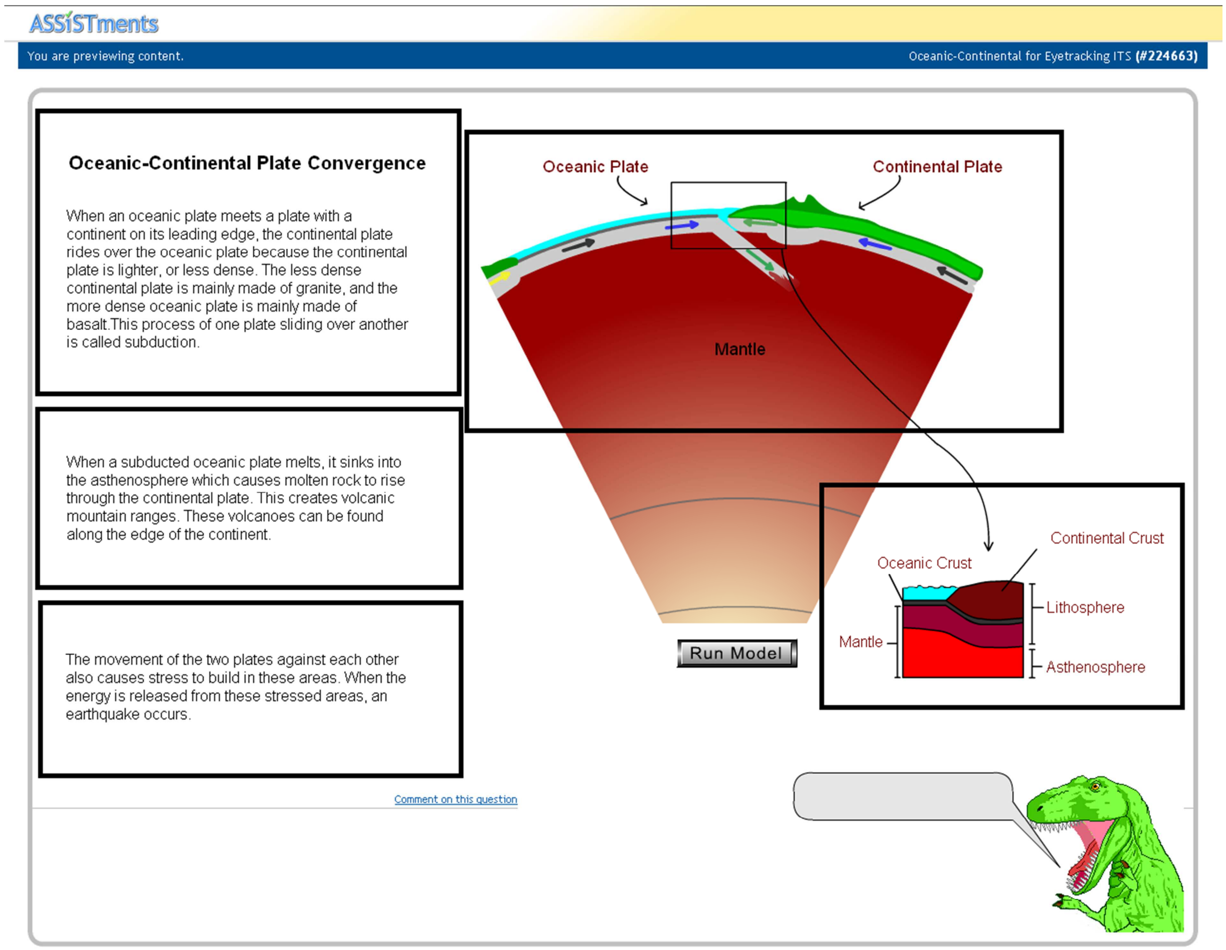


Figure 5 - Image of the oceanic-continental plate convergence microworld with defined regions

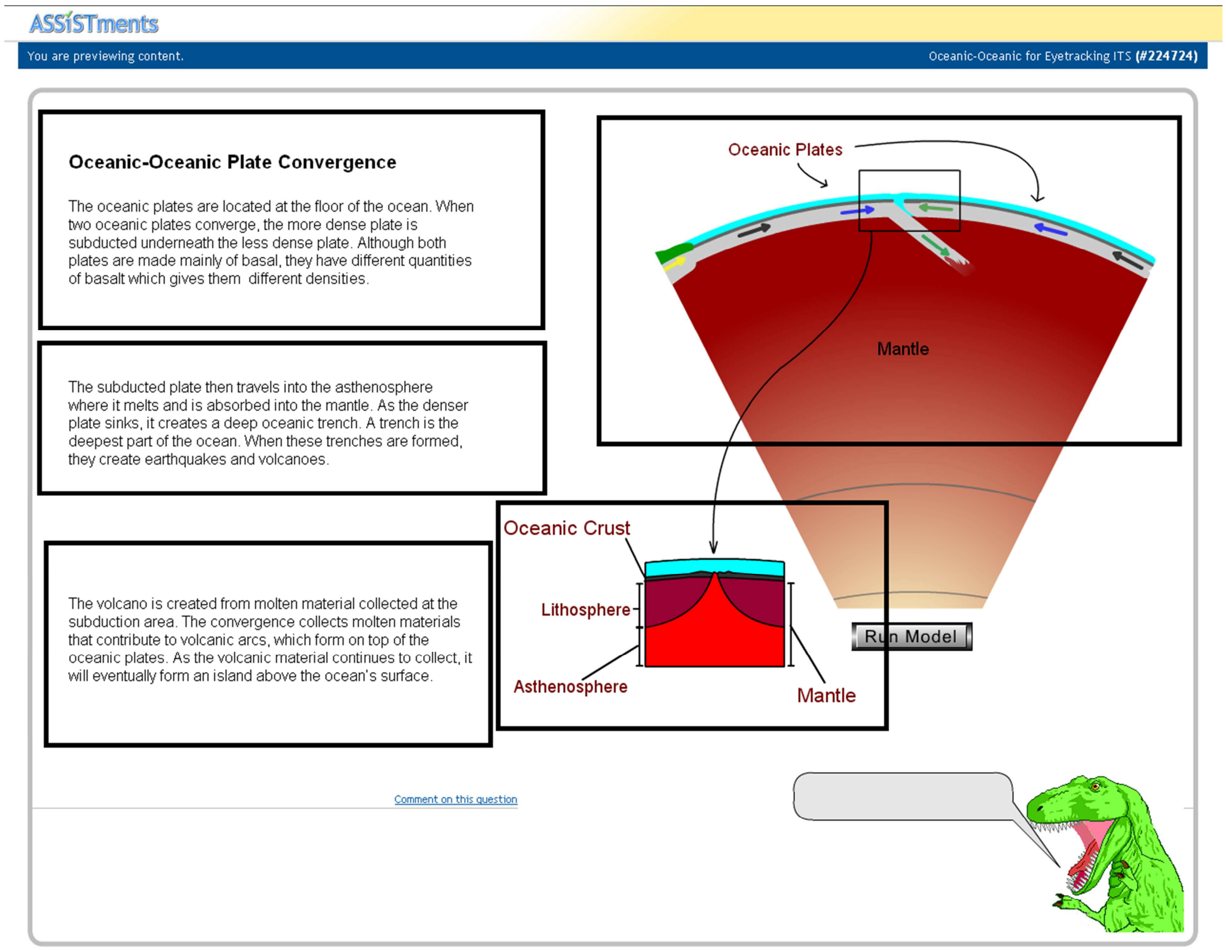


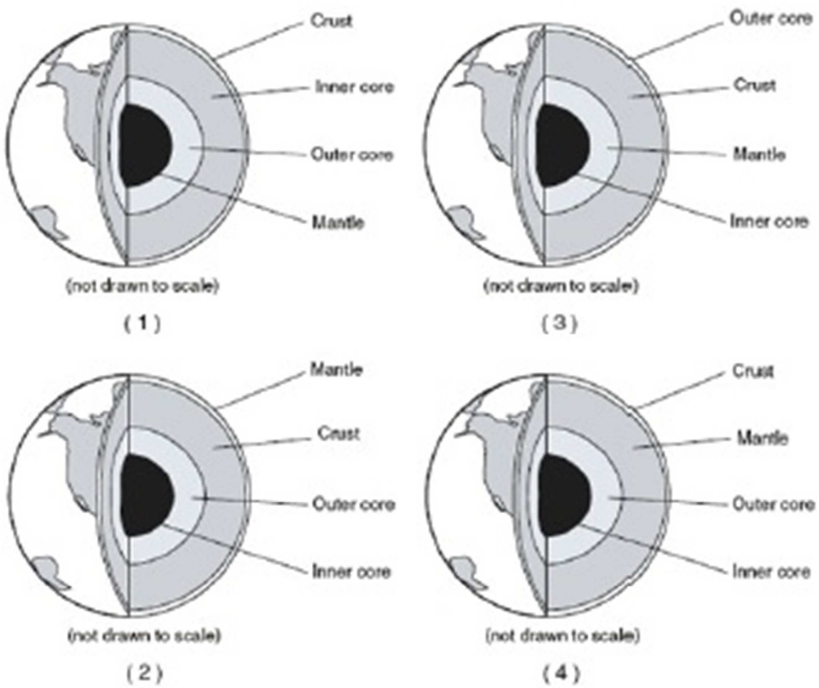
Figure 6 - Image of the oceanic-oceanic plate convergence microworld with defined regions

Appendix B - Pre/Post Test

Problem Set "PreITS" id:[38975]

1) Assistment #244042 "244042 - 237861 - Plate Tectonics IQP PRE Test. PRE is: 237860"

A) In which diagram are the layers of Earth correctly labeled?



- 1
- 2
- 3
- 4

B)

Of the following, which of Earth layers has the greatest density?

- crust
- mantle
- inner core
- outer core

C) The crust and upper mantle make up Earth's ________.

- lithosphere
- asthenosphere
- core
- continents

D)

The cycle of heating, rising, cooling and sinking is called a __________.

- subduction zone
- convergent boundary
- convection current
- conduction current

E) Active volcanoes on the Earth's surface are most likely to form at ________.

- transform boundaries
- the center of continents
- convergent oceanic-continental boundaries
- convergent oceanic-oceanic boundaries

F) Plates of the lithosphere float on the ___________.

- crust
- asthenosphere
- core
- atmosphere

G)

The heat from deep in Earth's interior is transferred to its crust by which of the following?

- conduction in the ocean
- convection in the mantle
- radiation from the solid core
- evaporation at mid-ocean ridges

H) Active volcanoes on the ocean's floor are most likely to form at ________.

- transform boundaries
- the center of continents
- convergent oceanic-continental boundaries

- convergent oceanic-continental boundaries
- convergent oceanic-oceanic boundaries

I)

Volcanic island arcs are associated with __________.

- transform plate boundaries
- divergent plate boundaries
- convergent oceanic-oceanic plate boundaries
- convergent oceanic-continental plate boundaries

J)

Partial melting and the production of magma takes place at _________.

- divergent plate boundaries
- convergent oceanic-oceanic plate boundaries
- convergent oceanic-continental plate boundaries
- all of these

K) Write a detailed explanation describing the different layers of the earth and the processes that happen inside the earth. Include all the information about these layers that you can so that a friend who did not do this activity could learn about it.

L) Write a detailed explanation of what happens in the different layers of the earth when mountains are forming. Include all the information you can so that a friend who did not do this activty could learn about it.

M) Write a detailed explanation of what happens in the different layers of the earth when volcanoes on the earth's surface are forming. Include all the information you can so that a friend who did not do this activty could learn about it.

N) Write a detailed explanation of what happens in the different layers of the earth when volcanoes on the ocean's floor are forming. Include all the information you can so that a friend who did not do this activty could learn about it.

Appendix C - Coding for Open Response Question 1, 2, 3, & 4

Static/Spatial components S1 (layers)	**Score**
earth	1
Has layers	1
crust	1
thin	1
Outermost layer	1
2 types of crust	1
Oceanic, continental	2
Oceanic under oceans	1
Continental underneath continents	1
mantle	1
Below crust	1
2900km/ 1802 miles thick	1
Uppermost part is solid	1
lithosphere	1
Made of crust and uppermost part of mantle	2
asthenosphere	1
Layer of mantle	1
soft	1
flowing	1
Rocky	1
Below lithosphere	1
Outer core	1
layer	
dense	1
hot	1
liquid	1
2190 km/ 1361 miles thick	1
Made up of iron, nickel	2
Inner core	1
dense	1
High pressure	1
solid	1
2680km/1665 miles thick	1
Made up of iron, nickel	2
Total spatial S1	
Causal/Dynamic components S1 (layers)	**Score**
Convention currents	1
Form circular flow of matter	1
Heat rises from core	2
Heat causes asthenosphere to circulate	2
Rises b/c it is less dense	2
Rises up through layer	1
When cool becomes more dense	2

When cool, sinks down	2
Total causal & dynamic S1	

Scoring for Question 2

Static/Spatial components S2 (c-c convergence)	**Score**
Plates with continents on them	1
Plate made up of same amount of land	1
Total spatial S2	
Causal/Dynamic components S2 (c-c convergence)	**Score**
Plate meet	1
Plates are equally dense	1
Neither can sink	1
Sink into mantle	1
Continents push together (instead of sinking)	1
Continents crush against the boundary	1
As plates collide, continents are crumpled/lifted	1
Produces highest mountain ranges	1
Plates are constantly pushing against one another	1
Plates are constantly straining against one another	1
Mountains get higher	1
Higher at a rate of ¼ inch per year	1
Total causal & dynamic S2	

Scoring for Question 3

Static/Spatial components S3 (c-o convergence)	**Score**
Oceanic Plate	1
Heavier, more dense than continental plate	1
Made of granite	1
Continental Plate	1
Lighter, less dense than oceanic plate	1
Made of basalt	1
Total spatial S3	
Causal/Dynamic components S3 (c-o convergence)	**Score**
Plates meet	1
Continental Plate slides over Oceanic Plate	1
One plate sliding over another plate = subduction	1
Subducted oceanic plate melts	1
Subducted oceanic plate sinks into asthenosphere	1
Sinking plate causes molten rock to rise	1
Molten rock rises through continental plate	1
Rising molten rock causes volcanic mountain ranges	1
Volcanoes found on edge of continent	1
Movement of plates against each other	1
Movement causes stress to build in these areas	1

Energy is released here	1
Earthquakes form as a result	1
Total causal & dynamic S3	

Scoring for Question 4

Static/Spatial components S4 (o-o convergence)	**Score**
Oceanic Plates (2)	2
Located on floor of ocean	1
Made of basalt	1
Plates are different densities	1
Due to different amounts of basalt	1
Total spatial S4	
Causal/Dynamic components S4 (o-o convergence)	**Score**
Plates converge	1
Denser plate slides over the less dense plate	1
One plate sliding over another plate = subduction	1
Subducted plate travels into asthenosphere	1
Subducted plate melts	1
Subducted plate is absorbed into mantle	1
As sinking plate sinks, creates oceanic trench	1
Oceanic trench is in the deepest part of ocean	1
Formation of trenches causes earthquakes	1
Formation of trenches causes volcanoes	1
Volcano is formed from molten material which collects at subduction area	1
Convergence collects molten materials	1
Molten materials contribute to volcanic arcs	1
Volcanic arcs form on top of oceanic plates	1
Volcanic materials collects, forms island	1
Island is located above ocean's surface	1
Total causal & dynamic S4	