The InVEST Volcanic Concept Survey: Exploring Student Understanding About Volcanoes

Thomas L. Parham Jr.[1], Cinzia Cervato[1], William A. Gallus Jr.[1], Michael Larsen[2], Jon Hobbs[1], Pete Stelling[3], Thomas Greenbowe[4], Tanya Gupta[4], John A. Knox[5], Thomas E. Gill[6]

ABSTRACT

Results from the Volcanic Concept Survey (VCS) indicated that many undergraduates do not fully understand volcanic systems and plate tectonics. During the 2006 academic year, a ten-item conceptual survey was distributed to undergraduate students enrolled in Earth science courses at five U.S. colleges and universities. A trained team of graders scored 672 completed surveys, coding responses to each item with a score, out of 3, based on accuracy and comprehensiveness. Questions requiring only basic content knowledge (e.g., terminology, volcano topology) received more high scoring responses than questions requiring higher thinking and deeper conceptual connections (association with plate tectonics, prediction of hazards and impacts on the environment). The mechanics of eruptions also appeared to be poorly understood. Special attention was paid to students' alternate conceptions about where volcanoes are likely to form. Male students, students highly interested in science, and students who lived in a volcanically active area received significantly higher total scores than other student groups. Science, technology, engineering, and mathematics (STEM) majors also performed significantly better than non-STEM majors. Understanding the nature of student comprehension and misconception may be useful for geoscience educators seeking to address student preconceptions and promote conceptual change.

INTRODUCTION

Many incoming college students have a weaker grasp of key geoscience concepts than would be expected if their secondary education had effectively addressed the National Science Education Standards. The fundamental theory of Plate Tectonics and several associated phenomena are included in the National Science Education Standards as seen earlier in the paragraph (National Research Council, 1996). Yet, recent studies (e.g., Libarkin and Anderson, 2005; Marques and Thompson, 1997) demonstrate that significant alternate conceptions related to plate tectonics persisted even at the undergraduate level.   Students also have a poor understanding of tectonically driven phenomena, such as earthquakes (Barrow and Haskins, 1996) and are confused about mountain building processes (Chang and Barufaldi, 1999; Muthukrishna et al. 1993). Studies of student preconceptions about volcanoes have been conducted on Italian high-school age children (Bezzi and Happs, 1994), but U.S. undergraduates' ideas about volcanoes remain largely unexplored.

Studies of undergraduate preconceptions are critical to the advancement of geoscience education in both secondary and tertiary settings. Concept retention is a significant concern for many secondary geoscience educators, but is often difficult to assess as students move on to work, training, or higher education. By exploring the understanding of undergraduates in large sample populations, we may offer secondary geoscience educators some perspective on the lasting effects of their efforts, at least among college-bound students. Post-secondary geoscience educators should be equally concerned with characterizing the nature of incoming students' prior knowledge, as this perspective is crucial to providing individualized instruction. Moreover, student preconceptions that are inaccurate can (and should) be specifically confronted with their scientific alternatives (Libarkin and Kurdziel, 2001). Education research has long suggested that this approach is effective in achieving conceptual change (e.g., Driver and Odham, 1986).

One proven means to assess students' prior knowledge is the administration of a concept inventory or conceptual knowledge survey. The physics education literature has explored student preconceptions since the 1980s (Halloun and Hestenes, 1985), eventually leading to the development of a famous instrument for assessing student knowledge, the Force Concept Inventory (Hestenes et al., 1992). Similarly, the Geoscience Concept Inventory (Libarkin and Anderson, 2006) represents a robust and highly successful instrument that addresses a broad variety of key geoscience concepts. Other geoscience concept test and questionnaire studies have taken a similarly broad approach (e.g., McConnell et al., 2006; Cervato et al., 2007), while some have focused specifically on a concept of particular concern, such as geologic time (Parham et al., 2005; Libarkin et al., 2007). In academic vernacular, these approaches could be distinguished as "conceptually extensive" and "conceptually intensive", respectively. This study follows the latter approach, employing a survey instrument designed specifically to explore student preconceptions about volcanic systems and eruptive processes – a highly dynamic system that many students, and even professionals, struggle to fully conceptualize.

As part of a long-term effort to explore student understanding about volcanoes and develop new computer-based teaching tools, the Interactive Virtual Earth Science Teaching (InVEST) project team created the Volcanic Concept Survey (VCS), a concept-intensive instrument that was designed to explore baseline levels of

[1]Department of Geological and Atmospheric Sciences, Iowa State University, Ames IA 50010; tparham@iastate.edu
[2]Biostatistics Center, George Washington University, Rockville MD 50010
[3]Department of Geology, Western Washington University, Bellingham WA 20852
[4]Department of Chemistry, Iowa State University, Ames IA 50010
[5]Department of Geography, University of Georgia, Athens GA 30602
[6]Department of Geological Sciences, University of Texas at El Paso, El Paso TX 79968

undergraduate student understanding about volcanoes, without the aid of support materials (notes, textbooks, etc) at the outset of introductory geoscience courses. Here we describe the design, development, and dissemination of the VCS and report on the demographics of the survey population. We explain in detail the scoring procedure and a preliminary study of grader reliability. Finally, we present qualitative and limited quantitative results of the survey, including areas of best and least understanding, together with a discussion of how these findings may inform the teaching practice of geoscience educators in both secondary and post-secondary environments.

SURVEY INSTRUMENT

The InVEST VCS development team consisted of faculty members in geology, meteorology, and chemical education. A volcanologist and experts in science education were consulted on issues of content validity and two statisticians provided expertise on survey design. Moreover, a graduate student in chemical education and a senior-level undergraduate in geoscience education helped ensure that question wording would be intelligible to an undergraduate student population.

The final survey consisted of two primary components: a free-response survey and an attached demographic questionnaire. The free-response section consisted of ten open-ended questions on a variety of volcanic concepts. These are available online (http://www.chronos.org/resources/DemoVolcanoTemplate.pdf) and are provided in Table 1. Many concept inventories use a multiple-choice format or employ Likert scales. However, we chose to leave questions open-ended to give students the opportunity to demonstrate the full extent of their thinking and establish connections between volcanic concepts. For example, question four of the InVEST instrument (*Think about the location of volcanoes on land around the world. Is there a pattern to their location, and if so, what might control that pattern?*) attempts to explore the topics addressed by question 13 from version 2.1.1 of the GCI (Figure 1; Libarkin and Anderson, 2009). Questions on the final version of the VCS were carefully chosen to span a variety of concepts related to volcanism and assess understanding across many levels of Bloom's Taxonomy (Figure 2; Bloom, 1956). Specifically allotted "free space" at the end of the instrument provided students with an opportunity to ask questions, clarify their responses, or share any of their own ideas about volcanoes that they felt had not been addressed.

13. The following maps show the position of the Earth's continents and oceans. The ●'s on each map mark the locations where volcanic eruptions occur on land. Which map do you think most closely represents the places where these volcanoes are typically observed?

A. Mostly along the margins of the Pacific and Atlantic Oceans

B. Mostly along the margins of the Pacific Ocean

C. Mostly in warm climates

D. Mostly on continents

E. Mostly on islands

FIGURE 1. Sample GCI question exploring volcanic pattern. Reproduced from online GCI v2.1.1 (Libarkin and Anderson, 2009 - https://www.msu.edu/~libarkin/gci.html).

The attached demographic questionnaire collected data on student background including gender, age, major, and learning preferences. This was designed to aid in the statistical analysis of survey results and support future in-depth research on the influence of learning styles and demographic factors on students' conceptual frameworks.

TABLE 1: QUESTIONS USED IN THE VCS WITH AN EXAMPLE OF IDEALIZED STUDENT RESPONSE

#	Question	Idealized Student Response[1]
1	Are all volcanoes similarly shaped? If not, how many distinct shapes can be seen? (Please illustrate your ideas below)	Some volcanoes (shield) are wide, broad and shallow-sloped. Some volcanoes (composite or stratovolcanoes) are steep-sided and rise to a peak, like an overgrown anthill. Some volcanoes (calderas) are partly destroyed, and look like giant holes in the ground.
2	What is the difference between lava and magma?	Magma is the combination of liquid rock, crystals, and gas below the surface. Lava is the same thing as magma but it is on top of the surface (exposed to air and/or surface water).
3	Describe the composition of a typical volcano. In other words, if you could cut a volcano in half, what would you see on the inside?	Many layers inside, often alternating between lava flows and ash (this is more likely in a stratovolcano). The layers slope away from the center of the volcano. There will also be many dikes (intrusions/veins/filled cracks) oriented more or less vertically, especially toward the center of the volcano (this could be called the "throat" or more correctly called the "conduit").
4	Think about the location of volcanoes on land around the world. Is there a pattern to their location, and if so, what might control that pattern?	Most volcanoes occur in linear belts and are often near the coast. Many underwater volcanoes occur in linear belts that run along the middle of the ocean floor. These volcanoes are controlled by the movement of tectonic plates, either running into each other (subduction) or spreading apart (spreading center or mid-ocean ridge). Some "hot spot" volcanoes occur somewhat randomly, and these are caused by thin, pencil-shaped plumes of hot material in the Earth's interior (mantle plumes).
5	Why does a volcano erupt?	Bubbles of volcanic gas become highly pressurized, and if the pressure of the bubbles within the magma exceeds the pressure of the rocks surrounding the magma, the rocks break and release the over-pressurized bubbly fluid.
6	What controls how explosive a volcanic eruption will be?	How many bubbles there are (which depends on how much water is present in the magma), how thick/sticky (viscous) the magma is (this depends on how much silica is present in the magma and how hot the magma is.)
7	How does water in a volcanic system affect how explosive a volcanic eruption will be?	If there is more water vapor in a magma, there will be more bubbles and each bubble will have more water in it, creating higher pressures. So, more water = more explosive eruptions
8	Draw a picture of an erupting volcano and identify as many features as you can.	Should show at least the following: lava flow, pyroclastic flow, and ashfall/ ash cloud
9	Volcanic eruptions can create natural hazards beyond the eruption of lava and ash. In the left column, list hazards caused by erupted material. In the right column, identify hazards caused by the interaction of these materials with their surrounding environment. [Two columns provided]	Eruption Material Hazards: Ash fall (including big rocks), Pyroclastic flow, Lava flow, Volcanic gases, etc Environmental Hazards: Lahar (mud flow), Lightning, Floods (melting of glacier, jokulhaup), Tsunami, etc
10	As specifically as possible, describe how a volcano might affect the following people or groups of people in the region: a. A farmer living at the base of the volcano: b. A tourist lodge along a stream flowing down from the volcano c. A pilot of a 747 flying through an area above the volcano: d. A group of skiers on the side of the volcano:	A.) Plants would smother under ash; later crops would thrive in rich soil. Risk for pyroclastic flow, lahar if near a river, too much ash fall crushing his/her home. B.) Proximity to stream puts them at great risk for lahar, also pyroclastic flow, perhaps floods (though these are rare). Fish population could be affected by ash in water, or by decrease in pH due to acidic gases in water. C.) Ash can reduce visibility, scratch windows, stall jet engines, scour wings and reduce lift. D.) Gases released (especially CO_2) may accumulate in low pockets causing asphyxiation, or can burn if super-heated; melting of snow can create floods and/or lahars.

[1]For each item, if students approximated these responses, graders were to award the maximum score (3).

As diversity is a primary concern in the modern geosciences, we were particularly interested in exploring differences between the conceptual understanding of students with various ethnic or cultural identities. This approach may offer some perspective on the efficacy of earth science instruction under-represented groups are receiving in the secondary environment, or, at the very least, how effectively these students have been able to use their educational experiences to master key concepts.

STUDY POPULATION

During fall 2006 and spring 2007 semesters, five participating colleges and universities administered the VCS within the first week of class, prior to instruction about volcanoes and plate tectonics. Some instructors chose to offer extra credit for participating. The institutional review boards of all institutions approved the instrument during summer 2006 and allowed the use of student responses for research. A total of 672 students [Iowa State University (n = 432), University of Texas - El Paso (n = 103), University of Georgia (n = 72), Western Washington University (n = 27), and Fort Valley State University (n = 38)] signed a consent form allowing their responses to be used for research purposes. Allotted time to complete the entire instrument (demographics and questionnaire) varied between twenty to thirty minutes.

The selection of participating school was guided by our interest in covering a broad and diverse student population and to include students from predominantly undergraduate institutions as well as research-extensive universities. The large proportion of students from Iowa State University (ISU), the project's home institution, is primarily due to the high enrollment levels (500+ students) in Iowa State's introductory physical geology each semester. Smaller samples collected from other institutions reflect both class size at the respective school

and, to a lesser extent, willingness of students to participate in the study. Overall, the survey population included 357 (53.1%) female and 315 (46.9%) male students. Students who identified themselves as members of racial/ethnic minority groups accounted for 30.4% of the total population.

The Fort Valley State population spanned two courses: one for science majors and one for non-science majors. Other courses were general education and large lecture-format "service" courses. Overall, open or undeclared and non-science majors (journalism, accounting, design, etc) represented 70% percent of the surveyed population. Even among these non-science majors, most had taken at least once earth science course prior to college. Of undeclared and non-science majors surveyed, only 9% had never had an earth science course, 20% had their most recent earth science course in middle school (grades 6-8), and 71% had taken an earth science course in high school (grades 9+). Science, technology, engineering and mathematics (STEM) majors tended to have taken earth science courses somewhat more recently, with only 5% having had no coursework, 17% having their most recent course in middle school, and 78% having had a course in high school. Regardless of major, roughly 43% of students had taken their most recent earth science course in 9th grade.

EVALUATION

Due to the open-ended nature of the questions, each survey needed to be reviewed and scored to allow any quantitative analysis. Moreover, the large number of responses necessitated delegation of scoring responsibilities among a group of graders, which consisted of three undergraduates, one graduate student, and three faculty members at Iowa State University. All graders attended a training session with members of the

TABLE 2. EXAMPLES OF STUDENT RESPONSES ACROSS SCORE LEVELS

Question	Strong Response[1] (Score Level 3)	Moderate (Score Level 2)	Weak Response (Score Level 1)
#3: Describe the composition of a typical volcano. In other words, if you could cut a volcano in half, what would you see on the inside?	"Layers of rock with lots of cracks on the sides and a deep chamber with magma coming up through the center."	"Rocks on the outside, magma inside" "Rocky cone and magma pipe inside"	"Magma at the bottom" "Lava tunnel inside" "Layers of magma"
#5: Why does a volcano erupt?	"Hot gases build up under pressure until magma breaks through the rock and escapes …"	"Lava is squeezed by pressure" "Too much heat and pressure inside"	"Magma overflows" "A build up inside as lava rises from the center of the earth" "Magma gets hot and expands"

1. Higher scores represent closer approximation of the idealized response (Table 1). For example, in the case of question 5, high-scoring responses discussed the effects of gas pressure on magma, while low-scoring responses were likely to neglect the role of volcanic gases and/or propose entirely different driving mechanisms, many of which indicate non-scientific preconceptions about Earth's interior.

development team, during which the idealized response for each question (Table 1) was shared, discussed, and modified if necessary. Then, each grader independently coded the same random selection of twenty surveys, assigning individual item responses a score between 0 and 3. During this phase, graders had the opportunity to discuss problems that arose during their reading of the responses, but each grader completed the scoring independently.

Non-informative responses ("I don't care", etc), or a failure to convey any measure of understanding received a score of zero. Often, these zero-level responses consisted of a single word unrelated to the question at hand, or were entirely blank. A score of 1 corresponded to a minimal level of understanding, while a score of 2 indicated further developed and/or accurate responses. Graders identified conceptual mastery, which was coded as a maximum score of 3, when students approximated the core ideas contained in the idealized response (Table 1). Graders were trained to de-emphasize terminology in favor of conceptual accuracy. However, some jargon-related issues did arise, and will be discussed later. Table 2 provides examples of high-scoring, moderate, and low-level responses to two VCS questions. Similar score coding approaches have been used to categorize and statistically analyze open-ended responses on highly vetted assessments of student knowledge, including the Trends in International Mathematics and Science Study (TIMSS) (Gonzales et al., 2008) and National Assessment of Educational Progress (NAEP) (National Assessment Governing Board, 2008).

To test for consistency and agreement among graders, as well as ascertain whether grader training had been effective, we performed statistical analysis on the scores of the twenty sample surveys utilizing intraclass correlations (ICC) with random grader effects (Shrout and Fleiss, 1979). A high ICC value implies that scores assigned by all graders to a given response were highly correlated. Thus, if ICC value is high (> 0.5), a response that was scored as a 2 by one grader was likely to be a given a 2 by any other grader. The higher the ICC value, the more agreement existed among all graders.

The results of this consistency test (Table 3) indicate that grader agreement was quite high for both total survey score and average item score. More importantly, grader agreement was strong for all but one of the items on the

TABLE 3: INTRACLASS CORRELATIONS (ICC)

Item	ICC	Item	ICC
Q1	0.700	Q9	0.596
Q2	0.818	Q10A[1]	0.394
Q3	0.676	Q10B	0.422
Q4	0.665	Q10C	0.542
Q5	0.708	Q10D	0.213
Q6	0.669	10Tot	0.526
Q7	0.665	Total[2]	0.857
Q8	0.603	Average[3]	0.857

[1]Parts of question ten (Table 1 #10a-d) were examined individually (10A, 10B, etc), but also factored into a composite question ten score (10Tot)
[2]Complete survey (Total)
[3]Average scores; Higher values indicate greater reliability

survey. Question 10 (Table 1) had less than ideal grader agreement on parts A, B, and D. However, when taken as a whole, grader agreement on question 10 (10Tot, Table 3) was much stronger. Thus, while graders occasionally scored individual parts of question 10 differently, the total number of points assigned for the entire question was consistent.

To further explore the possibility of grader effects, we constructed side-by-side boxplots for all graders with respect to total score distribution (Figure 2). Only minor variations in mean location and total variability were present, so no statistical corrections for grader effects were deemed necessary in further analyses. Upon conclusion of reliability testing, each grader scored a random subset of completed surveys, and compiled score data.

RESULTS

Survey results suggest that student understanding of volcanic processes was rather limited. The average total score on the instrument was twenty-five out of a possible thirty-nine points (64%). However, for the purpose of exploring students' understanding of specific concepts, individual question scores are more revealing. Generally, low-scoring questions were those requiring higher-thinking skills to analyze patterns or apply knowledge to make predictions (Figure 3). No student approximated the ideal response in three questions (9, 10B, and 10C). A further six of the ten questions (3, 5, 6, 8, 10A, and 10D) saw less than 1% of responses at the highest level (Score 3). Question 2 received the greatest relative proportion of high scores while question 8 was dominated by a large number of low scores.

Question 4 proved particularly interesting, as it addressed the locations of volcanoes around the world and asked students to think about what might be controlling their distribution. It should be noted that students were not provided with any visual aid (maps, diagrams, etc), but rather were expected to construct their
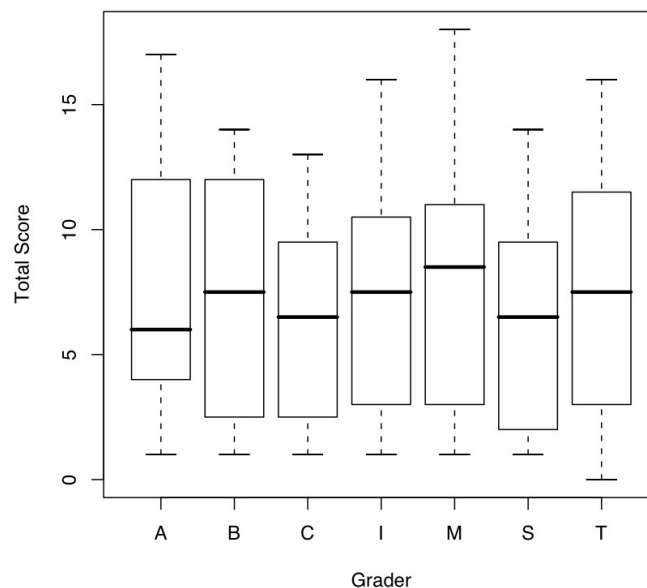

FIGURE 2. Boxplots for total score by grader. Each single letter ID represents a unique grader.

own conceptual imagery. While exposure to global volcanic, earthquake, and tectonic maps may promote connection between these phenomena, the goal of this survey was to assess base levels of prior student knowledge at the beginning of introductory geology courses without support material. In total, 512 students offered an answer. Of these, roughly half (n=258) correctly responded that there was indeed a pattern in the global distribution of volcanoes and, furthermore, indicated that this pattern was related to tectonic activity. The following are random examples of high-scoring responses:

" ... most are located where tectonic plates meet, but a few are located on hotspots"
"... around edges or the hot spots of the tectonic plates with high volcanic activity"
"Ring 'O Fire! Volcanoes often pop up at tectonic plate boundaries and hot spots"

In contrast, about the same number of students (n=254) failed to recognize a global distribution pattern of volcanoes and/or accurately describe the mechanism (tectonics) in control of that pattern (Table 4). Several types of preconceptions were apparent, the most predominant being a connection of volcanoes with nearby bodies of water and/or the belief that all volcanoes form on islands. This group accounted for nearly 17% of the total responses. The second most common preconception (15.2%) associated volcanism with "hot" or "tropical" climates, typically near the equator. Over 6% of the responses indicated that students believe volcanoes form due to "rough," "rocky," or "mountainous" terrain. Finally, 11% of students stated that there was no pattern or that volcanic formation was entirely random.

At a broader scope, analyses utilizing total score as an index of understanding found that male students (Mean Score = 7.838) performed better on the VCS than female students (Mean Score = 6.090). Caucasian students overwhelmed the population and, together with students in the "other" category (those who either marked the provided "other" option or declined to state), accounted for most of the highest scores. Small sample sizes complicate the interpretation of scores among most minority groups, but it appears that Native American students also score highly. The overall effects of ethnicity on total score are summarized in Figure 4. Geographical location also appears to be a significant factor: students from Western Washington University,
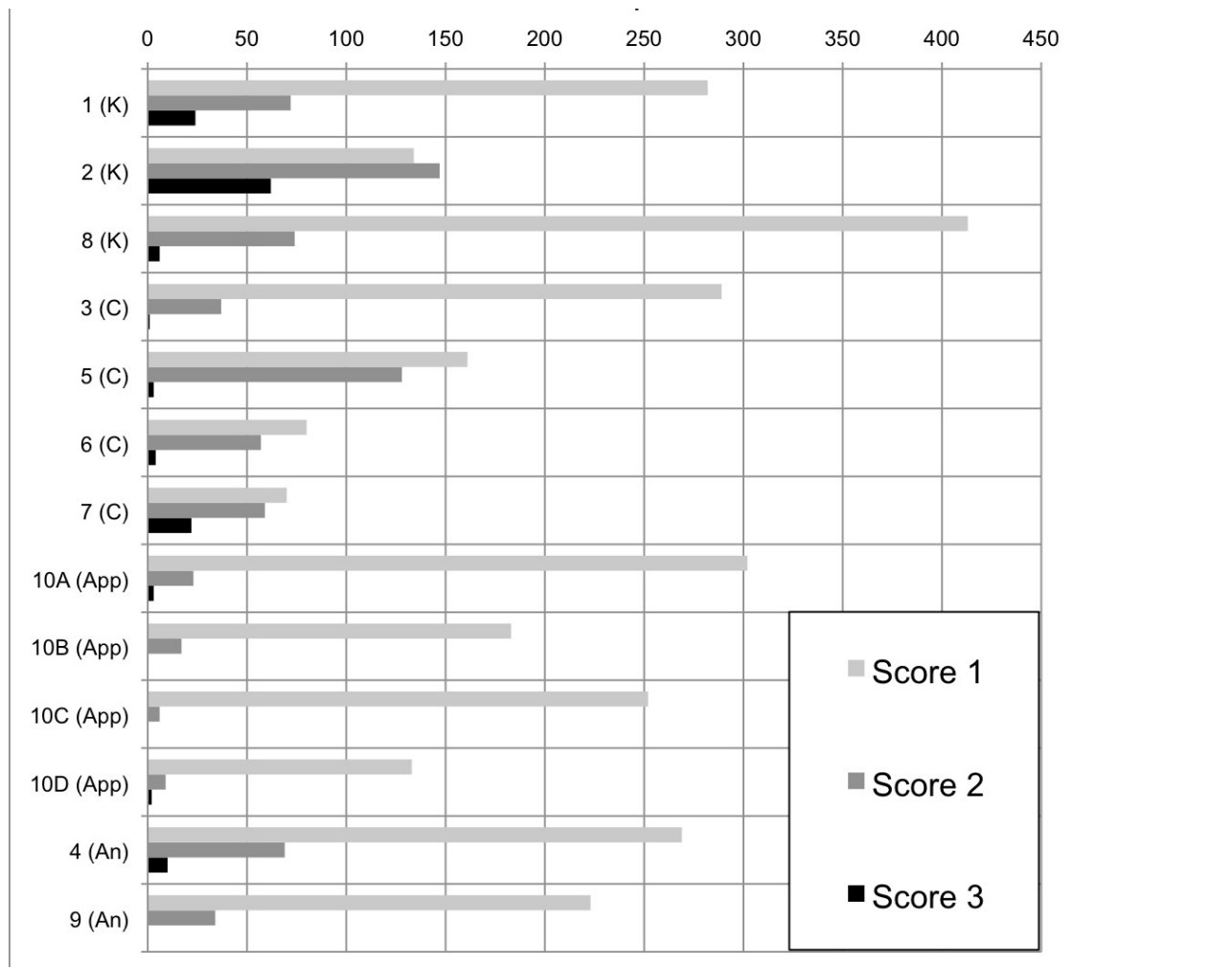

FIGURE 3: Distribution of scored responses to each VCS question, grouped by Bloom's Taxonomy: knowledge = K, comprehension = C, analysis = An, application = App. High-scoring responses are most common in the lower levels of the cognitive domain.

TABLE 4: PREVALENCE OF PRECONCEPTIONS REGARDING GLOBAL PATTERN OF VOLCANOES (QUESTION 4, TABLE 1)

Pop = 512	Limited to Nearby H_2O	Climate Controlled	Terrain Controlled	Entirely Random	Total
n	87	78	33	56	254
Sub %[1]	34.3	30.7	12.9	22.1	100
Pop %[2]	17.0	15.2	6.5	10.9	49.6

[1]Subgroup proportion (Sub%) is relative to the subgroup of responses that held some form of misconception
[2]Pop % is relative to all 512 responses to Question 4

the only participating school in a volcanically active area, performed much better than those from other schools (Figure 5). Students who claimed to be very interested in science did much better (Mean Score = 7.466) than those who were not (Mean Score = 5.365).

A t-test comparing the general education course and the course for science majors at Fort Valley State University showed that the difference in total score was significant (p = 0.022), with science majors performing better. Analysis of the entire data set confirmed this trend (Figure 6). Science Technology Engineering and Math (STEM) majors (including Engineering/Technology, Life Science, Natural Science, Mathematics, and Physical Sciences) perform significantly better than non-STEM majors (p < 0.0001). On average, physical and natural science majors received higher scores than other groups. In contrast, education majors received relatively low scores.

DISCUSSION

We begin our discussion of the survey results by discussing how Bloom's Taxonomy informed our interpretations. Groups of questions shall then be discussed in terms of the areas of greatest understanding (those in which students scored highly) and least understanding (where scores were notably lower). Potential complicating factors related to the design of survey questions will be addressed as they arise during this discussion. Question 4, which asks students to think about the global distribution of volcanoes and, ideally, to relate it to plate tectonics, is considered significant enough to merit independent examination. Finally, we shall explore some of the more interesting trends in total score across students' ethnic identities, lived geographies, and major of study. We will conclude by reflecting on the ways these results might be of use to geoscience educators in both secondary and post-secondary environments.
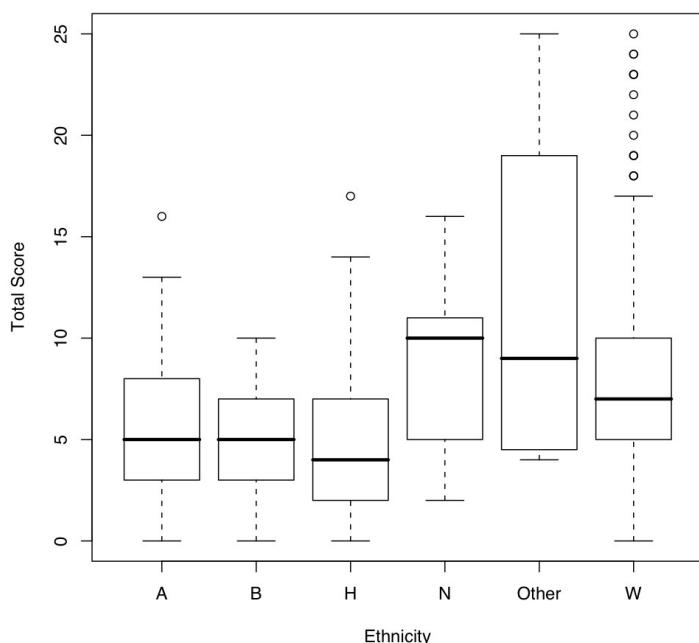

FIGURE 4. Boxplots showing the relative distribution of total scores by racial identity/ethnicity. Other = No response (n=12), A = Asian/Pacific Islander (n=34), B = African American/Black (n=55), H = Hispanic/Latino/a (n = 102), N = Native American/Alaskan Native (n = 9), W = Caucasian/White (n = 456). Unfilled circles represent "normal outliers" whose scores were at least 1.5 times the intraquartile range, or score distance between the high (Q3) and low (Q1) ends of the box, above the mean.
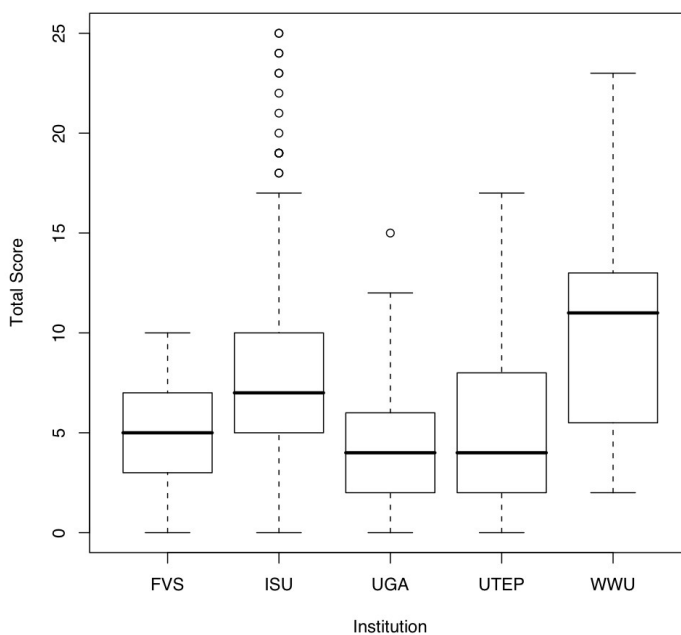

FIGURE 5. Boxplots showing the relative distribution of total scores by home institution. FVS = Fort Valley State (n = 38), ISU = Iowa State University (n = 432), UGA = University of Georgia (n = 72), UTEP = University of Texas - El Paso (n = 103), WWU = Western Washington University (n = 27).

Taxonomic Distinctions

The conceptual objective(s) of each VCS question may be ranked in Bloom's Taxonomy (Figure 2; Bloom, 1956). This helps separate questions that could have been effectively answered with only basic content knowledge from those that required higher-thinking skills such as the application of knowledge to analyze patterns or prediction of hypothetical results. Our results suggest that distinctions between Bloom's taxonomic levels are associated with score variation between questions. For example, both question 2 and 3 (Table 1) obtained a similar proportion of non-zero scored responses. However, question 2 obtained a significant portion of high scoring responses while question 3, which required some measure of higher thinking, saw many more low scores. Question 2, designed to address basic knowledge about what differentiates lava and magma and ranking at the lowest level of Bloom's Taxonomy, saw the overall highest proportion of Level 3 responses suggesting a general mastery of this concept. In contrast, question 3 assessed the understanding of the interior structure of a volcano as seen in cross-section. Though ranked just one level higher in Bloom's Taxonomy than the objective of question 2, this concept appears much less well understood. A similar trend can be observed throughout the instrument.

Areas of Greatest Understanding

Students tended to score more highly on questions requiring only basic levels of content knowledge or comprehension. Specifically, questions 1, 2, 7 and 8 saw the greatest abundance of high scoring responses. Strong
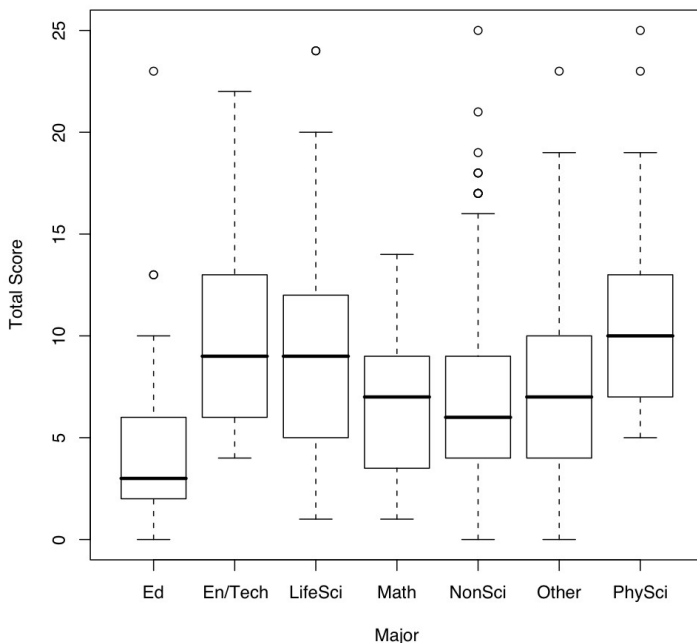

FIGURE 6. Boxplots showing the relative distribution of total scores by major. Ed = Education (n = 69), En/Tech = Engineering and/or Technology (n = 24), LifeSci = Biology (n = 51), Math = Mathematics/ Statistics (n = 24), Non-Sci = Humanities or other Non-Science (n = 434), Other = No option selected (n = 45), PhySci = Physical or Natural Sciences (n = 25).

performance on question #1 indicates that students seem to have a strong grasp of volcano topology, and that not all volcanoes are similarly shaped. High scores on questions 2 and 8 demonstrate familiarity with key terms such as lava and magma, as well as the ability to identify multiple features present during eruption. The high-scoring responses to question 7 are particularly interesting, as students demonstrate accurate knowledge that water vapor can increase the explosive character of an eruption, but yet the generally low performance on question 6 indicates that they often do not understand any other factors affecting explosiveness (magma viscosity, temperature, etc). It is possible that some students were, in essence, guessing in their response to question 7, assuming that more water in the system could generate a more explosive eruption simply because the survey inquired about water directly.

Areas of Least Understanding

Items that consistently obtained low scores include questions 3, 5, 6, 9, and 10 (Table 1). Question 3 asked students to draw a three-dimensional cross section of a volcano. Low scores on this item indicate that most students did not understand the inner workings of a volcano well enough to represent them graphically and may indicate a difficulty with 3-D spatial thinking. Question 5 saw a large number of non-zero responses, but a distinct lack of mastery-level understanding. Only one student explicitly identified gas pressure as the driving mechanism during a volcanic eruption. Roughly 32% of responses instead cited seismic activity. This may stem from the fact that earthquakes often occur as precursors or consequences of volcanic eruptions, although they do not cause eruptions themselves. Over 8% implicated simple overflow of magma within the chamber - similar to a free-flowing tap - suggesting misconception about the structure of Earth's interior. Low scores on question 6 may indicate a misunderstanding of what controls the explosiveness of a volcanic eruption. Very few students mentioned silica content or magma viscosity and its correlation with eruptive style.

Two low-scoring questions (9 and 10) were ranked in the higher cognitive domain of Bloom's Taxonomy. Thus the higher understanding required may partially explain the tendency for students to score poorly on these items. Questions 9 and 10 saw high levels of non-response and/ or non-informative responses, which obtained no points. When combined with the fact that non-zero scores on these items were at the low end of the scoring scale, this suggests that both environmental impacts and effects of eruption on human endeavors are poorly understood. While it may be true that even experts are likely to struggle with understanding the full impact of volcanism on humans and the environment, the VCS questions sought to address direct environmental hazards and realistically predictable short-term impacts on human activities. Interpretations of these questions may be complicated somewhat by the fact the the idealized response for #9 and parts of #10 (specifically 10B), utilize geological terminology. The intended goal of both items was not to test the mastery of jargon, but rather to

measure whether students were able to predict the environmental consequences of a volcanic eruption. As such, these questions will likely be revised in future work, and the wording of their idealized responses may need to be simplified.

Question #4

Recognition of a pattern connecting volcanoes, earthquakes, and plate motions is so fundamental in modern geology that it is difficult to over-emphasize the importance of students making this connection. The relatively high scores on question 4 should not overshadow the significant misconceptions demonstrated by nearly 50% of student responses (Figure 4). The most prevalent misconception attributed a global volcanic pattern to nearby bodies of water. While it is true that volcanoes often form in linear belts inland from the coasts, responses in this category used language that emphasized the involvement of surface water as a control mechanism and/or strictly limited the occurrence of volcanoes to islands surrounded by ocean waters.

Education research has shown that viewing images and diagrams can stimulate the rapid development of mental models (Butcher and Kintsch, 2004), though the information students perceive is not always correct. Thus, it is possible that students have made an association between volcanoes and water based on the fact that many images of volcanoes in the U.S. media come from Hawaii, Montserrat, or other volcanic islands. Prior to this study, work with the Geoscience Concept Inventory (GCI) has uncovered the tendency of introductory level students to assume that volcanoes are more common near the equator (Libarkin and Anderson, 2006). While the results of the InVEST instrument suggest that association of volcanoes with water may be an even more common misconception, we also confirm the strong presence of a climate-centered misconception. One student exemplified this thought by commenting: "... *where it's hotter, that's where they thrive best.*" Again, this misunderstanding may be the result of the prevalence of tropical volcanism in the media.

Perhaps more significantly, over 10% of responses to question 4 explicitly stated that there was no pattern to volcanic activity or that the process was "random." Other students (n=49) simply cited the "Ring of Fire" without further elaboration, indicating a familiarity with an important term, but little or no association of this pattern with the process driving it. It is likely that they heard about this term in their prior coursework, but they should also have explored the connection between plate tectonics and the global pattern of volcanoes, as it is included in the National Science Education Standards (National Research Council, 1996). Either students have not been taught about this important conceptual connection in their prior coursework, or the idea has not proven sufficiently durable to remain a part of their mental model of the Earth.

Even many of the best responses to question 4 did not demonstrate complete conceptual understanding. Within the subset of learners who correctly indicated that tectonic forces and/or features control global volcanic patterns, 43 inaccurately cited "fault lines", "weak spots" or "where the land is thinner". Although these factors are important in local control of volcanism, they do not account for the global pattern.

Demographic Trends

Returning to the analysis of trends in total score, it is apparent that some demographic groups performed significantly better than their peers. For example, male students and highly interested students tended to score better on the VCS. Other studies have concluded that male students tend to have more positive attitudes toward science (Trankina, 1993) and be more interested in exploring scientific topics (Jones et al., 2000). Thus, gender and interest level are likely to be self-reinforcing contributors to higher scores among male students. This requires further study, but underscores a need for instructors to specifically focus on stimulating interest and engagement among all students, including women and ethnic or cultural minorities. Innovative approaches such as small-group collaboration, peer learning, hands-on exercises, and activities based on real-life experiences make science courses more attractive to all students (Rosser, 1993).

High performance among the Native American student population may also merit future study. However, interpretation of why these students scored highly is complicated by a very limited sample. Only nine of the 672 students identified themselves as Native American, and five of these students also identified themselves as Caucasian or White. This very small population was similar to the overall survey population in many ways, including a roughly 1:1 gender ratio, being composed largely of non-science majors (n=7, two declared majors in life sciences), and reporting moderate interest in both general science and earth science. The Native American group also covered a broad geographical range, with students from every participating school except Fort Valley State. What may set this group apart is that all of its members had taken an earth science course, and over half (n=5) had taken an earth science course in 10th grade or later. Thus, we speculate that the high performance of the nine Native American students results from having been exposed to geoscience content more often or at a higher level. We also speculate that the high number of outliers within the Caucasian population is a consequence of the very large sample size (over 450 students). Being an order of magnitude larger than many of the minority populations, it is not surprising that the Caucasian group would contain many more students who performed beyond the range of normal variability.

Perhaps more than ethnic identity, students' lived geography may be a major influence on their understanding of volcanic systems, as reflected in their total score on the VCS. Students from Western Washington University, the only participating school located in a volcanically active area, generally performed much better on the VCS than students studying in areas without active volcanoes, possibly because volcanism has the potential to impact their immediate surroundings and is a part of their daily life. The high performance of students from Iowa State, which is located in a distinctly

inactive portion of North America, is more difficult to explain. There is no measureable difference in the recency of earth science coursework between Iowa State students and those from other participating schools, and while it is tempting to assume that Iowan students receive higher quality secondary instruction than do students in other states, the available data do not provide a conclusive explanation for why Iowan students would score highly on the VCS.

Since students who do not live in a volcanically active area appear to be working from somewhat of a disadvantage, educators without in-field resources and the advantage of local context must enrich their courses in other ways. While there is no substitute for field work and personal experience, computer simulations (e.g. Discovery Channel, 2009) may prove particularly useful to instructors in areas without active volcanism, especially when coupled with physical demonstrations (Erdogan, 2005; Harpp et al., 2005), analytical exercises (Harpp and Sweeny, 2002), and/or reflective writing strategies (e.g. Burke et al., 2006).

Finally, the fact that STEM majors had a better understanding of volcanic systems than their peers is not surprising, particularly in the case of the physical science sub-group. Several of the physical science students were listed as geology majors at the time of the survey. What is interesting, however, is the very low performance among education majors. This may point to a systemic problem, wherein those training to become elementary and secondary earth science and/or general science instructors are working with limited understanding of geoscience content. Examination of these students' content knowledge at the end of their geoscience coursework is beyond the scope of this study, and it seems likely that their grasp of core content would improve substantially. However, these findings underscore the critical need to provide all students with content-rich constructivist learning experiences – particularly pre-service educators.

CONCLUSIONS

This study utilized the concept-intensive Volcanic Concept Survey instrument to explore the pre-instruction (i.e. "baseline") understanding of volcanic systems among 672 undergraduate students enrolled in entry-level geoscience courses at five schools across the United States. We have shown that many undergraduates have a very limited understanding of volcanic systems. The knowledge that they do possess is often complicated by misconceptions and misunderstandings about where volcanoes form, why and how they erupt, and the broad effects of eruption on Earth systems and human endeavors. Moreover, the link to plate tectonics often is not understood. While most students can answer basic questions about volcano shape, differentiate between lava and magma, and label features of an erupting volcano, many do not demonstrate a deep enough understanding of volcanic concepts to deal with higher-level cognitive tasks related to predicting volcanic hazards.

Demographic data collected on the survey population indicates that several factors are associated with a better understanding of volcanic systems and processes, as represented by total score on the VCS instrument. The tendency of male students to score highly is likely related to self-reported high levels of interest in science. Certain ethnic groups including Native American and Caucasian students scored higher than their peers, which may be related to differences in educational background or factors that were not captured by the survey instrument. Students who live near volcanoes score considerably better on the VCS than those who do not, though students from Iowa State, a Midwestern research-extensive university, seem to defy this trend. Many students with declared majors in science, technology, engineering, and/ or mathematics tend to have a respectable grasp of concepts related to and governing volcanism when they enter their first geoscience course, as compared to non-science majors. Education majors begin with especially limited understanding of volcanic systems. Providing non-scientists, pre-service educators in particular, with individualized instruction to address their alternate conceptions is recommended.

These findings may be enlightening for secondary geoscience educators in search of a long-term perspective on what students have retained from their secondary coursework by the time they enroll in college. While basic information such as topology and terminology appears to be retained rather well, many students are either unable to conceptualize eruptive processes and the connection to plate tectonics during their secondary education, or these ideas are lost to students over time. Knowing the misconceptions students still hold after secondary instruction may help geoscience educators to focus their efforts and maximize retention of key concepts.

We believe the results of this study are equally useful for post-secondary geoscience educators, particularly those teaching introductory-level courses with a volcanology component. Constructivist teaching requires that instructors address the prior knowledge of their students and build upon it, promoting conceptual change when necessary. Armed with the knowledge of students' baseline understandings and misconceptions presented here, instructors can improve their instruction by targeting inaccurate ideas that are likely to exist among their students. Moreover, by administering the VCS or a similar instrument multiple times throughout the semester as a formative assessment, instructors may gauge whether their students understand volcanism within a solid conceptual framework.

REFERENCES

Barrow, L., and Haskins, S., 1996, Earthquake Knowledge and Experiences of Introductory Geology Students: Journal of College Science Teaching, v. 26, no. 2, p. 143-146.

Bezzi, A., and Happs, J.C., 1994, Belief Systems as Barriers to Learning in Geological Education: Journal of Geological Education, v. 94, p. 134-140.

Bloom, B.S., 1956, Taxonomy of Educational Objectives. Vol. 1: Cognitive Domain: New York, McKay, 205 p.

Burke, K.A., Greenbowe, T.J., and Hand, B.M., 2006, Implementing the Science Writing Heuristic in the General Chemistry Laboratory: Journal of Chemical Education, v. 83, p. 1032-1038.

Butcher, K.R., and Kintsch, W., 2004, Learning with diagrams: Effects on Inference and Integration of Information, In: Blackwell, A., Marriott, K., and Shimojima, A., Editors, Diagrammatic Representation and inference: Third International Conference, p. 337.

Cervato, C., Rudd, J.A. II., and Wang, V.E., 2007, Diagnostic Testing of Introductory Geology Students: Journal of Geoscience Education, v. 55, p. 357-363.

Chang, C., and Barufaldi, J.P., 1999, The use of problem-solving-based instructional model in initiating change in students' achievement and alternative frameworks: International Journal of Science Education, v. 21, p. 373-388.

Discovery Channel, 2009, Volcano Explorer, http://dsc..com/convergence/pompeii/interactive/interactive.html (accessed 19 February, 2009).

Driver , R. and Odham, V., 1986, A constructivist approach to curriculum development in science: Studies in Science Education, v. 13, p. 105-122.

Erdogan, I., 2005, Controlled Volcanism in the Classroom: A Simulation, Science Activities: Classroom Projects and Curriculum Ideas, v. 42, p. 21-24.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., Brenwald, S., 2008, Highlights from TIMSS 2007: Mathematics and Science Achievement of U.S. Fourth- and Eight-Grade Students in an International Context, National Center for Education Statistics, 112 p.

Halloun, I.A., and Hestenes, D., 1985, The initial knowledge state of college physics students: American Journal of Physics, v. 53, p. 1043-1055.

Harpp, K.S., Koleszar, A.M., and Geist, D.J., 2005, Volcanoes in the Classroom: A Simulation of an Eruption Column: Journal of Geoscience Education, v. 53, p 173-175.

Harpp, K.S., and Sweeny, W.J., 2002, Simulating a volcanic crisis in the classroom: Journal of Geoscience Education, v. 50, p. 410-41.

Hestenes, D., Well, M., and Swackhamer, G., 1992, Force Concept Inventory: The Physics Teacher, v. 30, p. 141-158.

Jones, M. G., Howe, A., and Rua, M. J., 2000, Gender differences in students' experiences, interests, and attitudes toward science and scientists: Science Education, v. 84, p. 180-192.

Libarkin, J.C., and Anderson, S.W., 2009, The Geoscience Concept Inventory Version 2.1.1, https://www.msu.edu/~libarkin/gci.html (accessed 13 February, 2009).

Libarkin, J.C., Kurdziel, J.P., and Anderson, S.W., 2007, College student conceptions of geologic time and the disconnect between ordering and scaling: Journal of Geoscience Education, v. 55, p. 413-422.

Libarkin, J.C., and Anderson, S.W., 2006, The Geoscience Concept Inventory Version 1.0, http://newton.bhsu.edu/eps/gci.html (accessed 16 April, 2008).

Libarkin, J.C., and Anderson, S.W., 2005, Assessment of Learning in Entry-Level Geoscience Courses: Results from the Geoscience Concept Inventory: Journal of Geoscience Education, v. 53, p. 394-401.

Marques, L., and Thompson, D., 1997, Misconceptions and conceptual changes concerning continental drift and plate tectonics among Portugese students aged 16-17: Research in Science and Technological Education, v. 15, p. 195-222.

McConnell, D.A., Steer, D.N., Owens, K.D., Knott, J.R., Van Horne, S., Borowski, W., Dick, J., Foos, A., Malone, M., McGrew, H., Greer, L., Heaney, P.J., 2006, Using Conceptests to Assess and Improve Student Conceptual Understanding in Introductory Geoscience Courses: Journal of Geoscience Education, v. 54, p. 61-68.

Muthukrishna, N., Camine, D., Grossen, B., and Miller, S., 1993, Children's alternative frameworks: Should they be directly addressed in science instruction?: Journal of Research in Science Teaching, v. 30, p. 233-248.

National Research Council, 1996, National Science Education Standards: Washington D.C., National Academy Press, 272 p.

National Assessment Governing Board, 2008, Science Framework for the 2009 National Assessment of Educational Progress: Washington, D.C., US Department of Education, 155 p.

Parham, T., Cervato, C., Reed, J., Keane, C.M., Peart, L., Ross, M., Scotchmoor, J.G., Seber, D., Snyder, W.S., and Springer, D., 2005, The CHRONOS Online Questionnaire on Geologic Time and Earth History for 6-12 Grade Students and Teachers: A first step towards a succesful community-based E&O Program: Geological Society of America Abstracts with Programs, v. 37, no. 7, p. 148.

Rosser, S.V., 1993, Female-Friendly Science -- Including Women in Curricular Content and Pedagogy in Science: The Journal of General Education, v.42, n.3, p. 191-220.

Shrout, P.E., and Fleiss, J.L., 1979, Intraclass correlations: Uses in assessing rater reliability: Psychological Bulletin, v. 86, p. 420-428.

Trankina, M. L., 1993, Gender differences in attitudes toward science: Psychological Reports, v. 73, p. 123-130.

[i] http://www.chronos.org/resources/DemoVolcanoTemplate.pdf