Proceedings of the Indiana Academy of Science
(1996) Volume 105 p. 251-267

STUDENT EVALUATION OF AN UNDERGRADUATE EARTH SCIENCE CLASS AS A FUNCTION OF TOPIC DIFFICULTY AND TOPIC INTEREST

William John Boone
School of Education 3068
Indiana University
Bloomington, Indiana 47405

ABSTRACT: A survey of over 250 undergraduate students enrolled in an earth science class for non-majors revealed a wide range of attitudes with respect to 15 key topics presented during lecture and lab. The students provided attitudinal data on their interest in the various topics as well as on how difficult each topic was to understand. Whether or not a topic was perceived as interesting was not dependent on whether or not it was easy to understand. Class material which called for spatial visualization was considered quite difficult by the students.

KEYWORDS: Attitude, earth science, science education, undergraduate.

INTRODUCTION

Undergraduates often enroll in science courses simply to fulfill degree requirements. Commonly, these introductory courses are designed with the non-major in mind, and they provide non-majors with both lecture and laboratory experiences. In an effort to better understand the views of undergraduate non-majors on the topics covered in one earth science class, an attitudinal survey was administered to over 250 students. The survey asked the students to rate the lectures and labs as a function of topic interest and difficulty.

The attitudes of students should not be the sole factor shaping content in an earth science course; however, such data can be used to raise both the level of student learning and the quality of teaching. Because these courses are only a minor component in the student's course work, such classes must engender an interest in earth science (or science in general) which will last a lifetime.

The student perceptions described in this paper should be useful to all earth science teachers. Even if the geology courses at your institution differ greatly in content from the courses discussed here, the techniques outlined in this paper should serve as a useful guide for the construction of related surveys. The data presented here also represent a baseline against which further studies can be compared.

RATIONALE AND PREVIOUS WORK

Two factors that affect a course's success are the students' perceptions of topic difficulty and topic interest. If a topic is viewed as difficult to understand, then the instructor should bear this perception in mind when he/she teaches. Difficult topics could potentially frustrate the students and turn them off to the field. The topics viewed as less difficult may point to presentations using more successful teaching techniques that could be expanded to other parts of the course. Just as students perceptions of topic difficulty are useful in planning a course, knowledge of their interests can also be helpful. Using topics whose interest levels are highly rated, an instructor may be able to identify teaching techniques and subject material that would make other parts of the course more interesting to the students. Finally, a comparison of class topics ordered by both difficulty and interest can aid earth science instructors. Topics of low interest and great difficulty may need to be reorganized so that greater interest can be generated. When difficult topics are rated high in interest, a key teaching technique used for those topics might increase the appeal of other topics.

A number of researchers have developed and/or utilized attitudinal instruments to obtain information helpful to science teachers. Instruments have been designed and administered to collect attitudinal data from elementary school students (Davis, Lang, Taylor, Knight, and Sims, 1992), secondary school students (Towse, 1983; Renner, Abraham, and Birnie, 1985), teachers in training (Lucas and Dooley, 1982; Koballa, 1984), and active teachers (Lombard, 1982; Finley, Stewart, and Yarroch, 1982).

Some research has also been carried out on student attitudes and earth science topics. Fortner and Mayer (1991) conducted a longitudinal study of the attitudes and understanding of fifth and ninth graders on oceans and the Great Lakes, using an attitudinal instrument based on semantic differential items. Crawley and Coe (1990) studied the attitudes of middle school earth science students toward enrolling in high school science classes.

Studies involving college students in earth science courses are also found in the literature. Kern and Carpenter (1984) reported on the attitudes, values, and interests of undergraduates to 30 topics in an earth science course in which roughly half of the surveyed students were enrolled in a traditional earth science laboratory, while the others were enrolled in a field-oriented lab. Carpenter (1982) discussed attitudinal changes in undergraduates towards a physical geology and environmental earth science class over time. Buchwald and Bybee (1990) discussed the role of geologists in the reform of science education, emphasizing the need to develop positive attitudes about knowledge and inquiry in children.

Table 1 lists the topics which were used to assess attitudes among the undergraduates participating in this study. Although several studies have focused on how to teach a number of these topics, none appear to have involved the collection of attitudinal data. For example, Eves and Davis (1987) discussed teaching about igneous rock in the lab, while Birdd (1990) described activities that could be used to present information on the rock cycle as well as on the development of metamorphic rock. Activities on sedimentary rock formation were outlined by Perdue (1991), while mineral identification has been discussed by numerous authors (e.g., Pasteris, 1983). Happs (1982) compared the views of children and scientists in regard to glaciation; wind erosion activities were described by McLure (1991); shoreline depositional models have been discussed by Miall (1982); and Bart (1991) has outlined topographic map activities. Although college texts discuss relative age and cross-cutting, few studies have discussed how to improve the teaching of these topics in the classroom (Newman, 1983; Oringer, 1985). Many activities have been published on mineral resources (Sherman and Stone, 1992).

METHODOLOGY

Instrument Development. The survey required the students to rate 15 key topics (Table 1) by difficulty and interest using a four-step scale: very difficult, difficult, easy, and very easy. The selected topics covered almost all of the major weekly lecture and laboratory topics. Before the final survey was completed, a draft was returned to the course instructor for reaction and feedback. Following the instructor's review, the survey was evaluated by seven student judges, who had completed an earth science course and were familiar with the rating scales. The student judges were asked to evaluate the clarity of written directions, the ease with which the rating categories could be used, and the wording of each item. Most of the changes suggested by the judges were made in the survey format. An example of one change was in the visual format of the survey. The judges suggested that a different font be used for the rating category descriptors. The modified version of the instrument was then re-evaluated by the judges. The survey was revised and then piloted using ten students who had taken this course. Written comments were solicited from these ten students.

Data Collection. The final version of the survey was administered to students attending a large Midwestern university during the twelfth week of classes (three weeks prior to the end of the semester). Students in 24 laboratory sections were asked to respond to the survey. None of the sections contained more than 20 students. Over 250 students attended lab sessions during the week the surveys were distributed. Of these students, 95% completed the survey.

The text, lab manual, and instructor are important factors which affect students' attitudes toward class topics. To see how applicable the responses were to earth science courses in general, the survey was also administered to students attending a large state university in the southern United States. In general, the same type of student (non-major) enrolled in both courses, and both classes were organized in a similar manner. Not all of the topics covered at the Midwestern university were presented at the Southern university, but at both institutions, similar class formats and texts (Lutgens and Tarbuck, 1989; Tarbuck and Lutgens, 1993) were utilized. Fifty-six students at the Southern university completed the survey. Data from the Midwestern university are presented first followed by the data from the Southern university.

Data Evaluation. The stochastic Rasch model (Rasch, 1960) was used to evaluate these data. This method was used because ordinal attitudinal scales first must be converted to interval scales, especially if one wishes to use parametric tests to analyze the data. The necessity of conversion can best be understood

Table 1. The list of topics surveyed at a Midwestern university. Topics 11-15 were labs. One survey asked students to indicate their level of interest in each of the topics by choosing from four possible responses: 1) very disinteresting; 2) disinteresting; 3) interesting; or 4) very interesting. A second survey asked students to indicate how difficult each topic was to understand by selecting from four possible responses: 1) very difficult; 2) difficult; 3) easy; or 4) very easy. The topics presented to the students at a second institution are listed in Table 4.

1	The difference between rocks and minerals
2	The rock cycle
3	Using cross-cutting relationships to determine the relative age of rocks
4	The way sedimentary rocks are formed
5	The way igneous rocks are formed
6	The way metamorphic rocks are formed
7	Mineral resources in the world
8	The landforms caused by glaciers
9	The role wind plays in erosion
10	The way shorelines are created
11	How to identify minerals (lab 1)
12	How to identify igneous rocks (lab 2)
13	How to identify sedimentary rocks (lab 3)
14	How to identify metamorphic rocks (lab 4)
15	How to read a topographic map (lab 5)

by noting that a step in attitude from very easy to easy does not necessarily represent the same quantifiable change in attitude as a step from easy to difficult (Thurstone, 1925, 1926, 1928; Wright and Masters, 1982). This problem is often overlooked because false linearity of scales is often suggested by the computer coding of attitudinal data. The Rasch method was also selected because: 1) evaluation is possible even when the respondents do not answer every item; 2) measurement errors of survey items and respondents are reported; and 3) idiosyncratic responses by the students can be easily detected (Wright and Masters, 1982). Many researchers have utilized this method to solve a variety of problems connected with the calibration of test items (Kelley, 1979; Wongbundhit, 1985). Rasch measures have also been applied in science education (Preece, 1979; Monk, 1984; Meyer, 1989). The Bigsteps computer program (Wright and Linacre, 1991) was utilized to calibrate measures of both respondents and items.

When categories such as very easy, easy, difficult, and very difficult are coded, many evaluators assign weights to each response. A very easy might be named a 1, while an easy is considered a 2. A rating of difficult is assigned a 3, and a very difficult is called a 4. Naming categories with numbers is fine. However, one can not immediately use these numerical identifiers for statistical calculations. When the calculations are made, the implicit assumption is that a response of easy is equidistant from a view of very easy and difficult. Therefore,

Table 2. The surveyed topics ordered based on the interest ratings of the Midwestern university students. Items at the top of the table were viewed as more interesting than those listed at the bottom. Sample free statistics calculated for the survey indicate a reliability of 0.90. The raw averages were corrected for non-linearity, and each item's measurement error was calculated so that Figure 1 could be constructed and parametric tests utilized for analysis.

I	Topic	Count	Score	Raw Average	SD	Measure (Logits)	Error (Logits)
8	Glacial landforms	284	287	2.01	0.69	0.92	0.12
7	Mineral resources in the world	278	299	2.08	0.64	0.68	0.12
10	Way shorelines created	281	322	2.15	0.68	0.41	0.12
9	Role of wind in erosion	283	345	2.22	0.67	0.10	0.12
11	Identify minerals	283	350	2.24	0.80	0.01	0.12
13	Identify sedimentary rocks	283	357	2.26	0.78	-0.08	0.12
12	Identify igneous rocks	283	359	2.27	0.78	-0.11	0.12
14	Identify metamorphic rocks	282	360	2.28	0.78	-0.15	0.12
1	Difference between rocks and minerals	283	363	2.28	0.61	-0.17	0.12
6	Way metamorphic rocks form	285	366	2.28	0.67	-0.18	0.12
5	Way igneous rocks form	285	366	2.28	0.64	-0.18	0.12
3	Cross cutting for age	280	366	2.31	0.70	-0.26	0.12
15	Reading topographic maps	270	360	2.33	0.83	-0.28	0.12
4	Way sedimentary rocks form	284	378	2.33	0.64	-0.35	0.12
2	Rock cycle	284	378	2.33	0.68	-0.36	0.12

I = Survey item number as shown in Table 1.
Count = The number of respondents answering a survey item.
Score = The raw score calculated from all survey responses to a single item using the scale: 1 = very interesting; 2 = interesting; 3 = disinteresting; and 4 = very disinteresting.
Raw average = The arithmetic mean of the raw scores.
SD = The standard deviation of the raw score.
Measure = Item measure in logits.
Error = The standard error of the item measure in logits.

when the students complete this type of attitudinal survey, the numbers used to code the data imply an ordering of attitudes (5 is greater than 4, and 4 is greater than 3, etc.) but not a known spacing. Thorndike (1904) stressed the need to convert raw scores to measures, and Thurstone (1925, 1926) developed partial solutions to this measurement problem. However, not until the Rasch model was used by Wright and others did a technique for converting raw scores to measures become available. Item response theory can now be used to adjust non-linear rating scales (Andrich, 1982; Rost, 1988). The basic stochastic model used to convert raw scores of coded responses into true measures is discussed by Wright and Masters (1982), Rasch (1960), Andersen (1973, 1977), and Barndorff-Nielsen (1978).

Why use the Rasch model instead of factor analysis for these data? A complete discussion of this question would require a separate paper. Therefore, only a minimum number of theoretical and practical points are presented here. First, in factor analysis, ordinal scores are viewed as interval measures without error, while in Rasch analysis, ordinal responses are modeled as stochastic manifestations of linear parameters estimated with measurement error. Second, the use of the Rasch model enables irregular responses to be identified through fit statistics. With factor analysis, variance increases slightly as the result of irregular responses, but these irregular responses are not detectable from factors. Third, when factor analysis is utilized, the identification of solitary biased items is difficult. Bias eigenvalues are usually insignificant. The Rasch model, however, provides item fit statistics that allow poorly functioning items to be identified. Fourth, miskeyed items can cause problems in an analysis. In a factor analysis, eigenvalues which might highlight miskeys are usually insignificant. Misfit statistics from the Rasch model, however, can point to miskeyed items (see Wright (1991) for a discussion of the issues mentioned above). Finally, Duncan (1984, p. 206) made the following statement about the indeterminancy of factor analysis:

> The technology of factor analysis has to do with the estimation of the loadings from the observable correlations and the determination of the acceptability of the model, which is assessed by its ability to reproduce those correlations within the limits of sampling error. If the model is accepted, there is still an indeterminacy in the estimation of factor loadings inasmuch as systems of factors differing from those in $\rho_{12} = \sum_j \alpha_{1j} \alpha_{2j}$, but mathematically equivalent to them in the sense that they produce exactly the same correlations on the left-hand side (of the equation above), may be chosen in an infinity of ways.

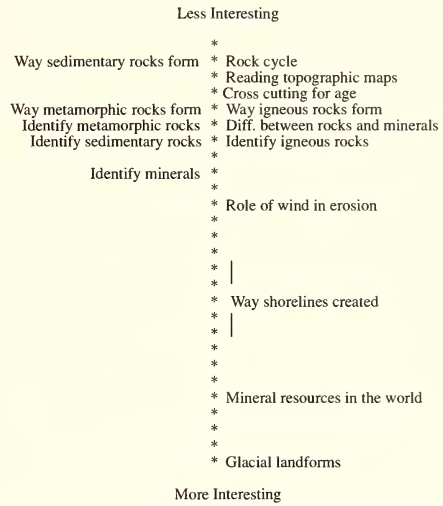
Discussions regarding the issues associated with factor analysis are provided by Reckase (1979), Abdullah (1989), and Raatz (1985).

The general relationship between logit values and raw scores can be seen by looking at the "Measure (Logits)" and "Raw Average" columns of Tables 2 and 3. The "Raw Average" column reports the average response of all survey takers to each item using the original coding. For the interest rating of topics, the coding scale was: 1 = very interesting; 2 = interesting; 3 = disinteresting; and 4 = very disinteresting. For the difficulty rating of topics, the coding scale was: 1 = very easy; 2 = easy; 3 = difficult; and 4 = very difficult. Thus, as Table 2 shows, the study of glacial landforms was rated on average as a 2.01 (on the 1-4 non-linear interest scale) by the students, indicating that, on average, the students not only found this topic interesting, but, of the surveyed topics, it was the most interesting. The difficulty of the same topic was rated as a 2.36 (on the 1-4 non-linear difficulty scale) by the students (Table 3), nearly halfway between easy and difficult, but, of the surveyed topics, the study of glacial landforms was viewed as one of the most difficult.

RESULTS

Data from the Midwestern University. The results of data analysis are shown in Tables 2 and 3 and Figures 1 and 2. In Figure 1, the item at the base of the thermometer, glacial landforms, was the most interesting class topic. The two items at the top of Figure 1, the way sedimentary rocks form and the rock cycle, were the least interesting class topics. Moving from the bottom of Figure 1 to its top, the items are rated according to decreasing interest. Figure 2 shows the survey items ranked on the difficulty scale. In this Figure, reading topographic maps appears at the top of the Figure, indicating that this topic was the most difficult, while the topic listed at the bottom of the Figure, the role of wind in erosion, was rated the least difficult.

Figure 1. A plot of the 15 Midwestern university class components as a function of item rating displayed on a linear scale corrected with the stochastic model. The standard error of all items was approximately that shown by the error bars for the topic, the way shorelines are created. Those items rated as more interesting are listed at the bottom of the chart, while those rated as less interesting are at the top. The full text of all items is given in Table 1.



Implications of Topics Ordered by Interest at the Midwestern University. Three topics were viewed as quite interesting: 1) the study of glacial landforms; 2) mineral resources; and 3) the creation of shorelines (Table 2 and Figure 1). One reason for their high rating might be that students feel familiar with these topics. Most students have seen pictures of mountains with alpine glaciers and have walked along a shoreline of some sort. Much of the geographic region from which this university attracts its students was affected by continental glaciation. Thus, many students are exposed to these concepts at some point in their precollege education, even though they may not have taken an earth science course. The mineral resources of the world (i.e., searching for gold, silver, oil, etc.) might be interesting because students are generally familiar with mining, panning for gold, and oil exploration.

Four topics were clearly less interesting: 1) sedimentary rock formation; 2) determining the relative age of a rock; 3) the rock cycle; and 4) using topographic maps. The low interest level generated by age determination and using topographic maps might be related to the need to mentally translate a two-dimensional image into a three-dimensional one. Lectures and labs involving cross cutting usually ask the students to interpret a cross section in which a variety of igneous intrusions are present in order to determine the sequence in which geologic events took place. Their interpretation is further complicated by a fourth variable, time, which makes translating images from paper into three dimensions more difficult. Students may view topographic map reading as having no future use, or they may have difficulty in translating two-dimensional maps into

Table 3. The surveyed topics ordered based on the difficulty ratings of the Midwestern university students. Items at the top of the table were viewed as less difficult than those listed at the bottom. The sample free statistics calculated for the survey indicate a reliability of 0.93. The raw averages were corrected for non-linearity, and each item's measurement error was calculated so that Figure 2 could be constructed and parametric tests utilized for analysis.

I	Topic	Count	Score	Raw Average	SD	Measure (Logits)	Error (Logits)
9	Role of wind in erosion	282	312	2.11	0.53	0.56	0.12
6	Way metamorphic rocks form	283	323	2.14	0.55	0.41	0.12
4	Way sedimentary rocks form	282	323	2.15	0.52	0.40	0.12
1	Difference between rocks and minerals	280	324	2.16	0.63	0.34	0.12
5	Way igneous rocks form	281	327	2.16	0.51	0.33	0.12
7	Mineral resources in the world	280	329	2.18	0.54	0.27	0.12
2	Rock cycle	281	335	2.19	0.60	0.21	0.12
12	Identify igneous rocks	279	336	2.20	0.66	0.16	0.12
14	Identify metamorphic rocks	281	341	2.21	0.69	0.12	0.12
13	Identify sedimentary rocks	279	347	2.24	0.67	0.01	0.12
11	Identify minerals	277	355	2.28	0.72	-0.13	0.12
10	Way shorelines created	273	366	2.36	0.65	-0.45	0.11
3	Cross cutting for age	282	405	2.44	0.74	-0.71	0.11
15	Reading topographic maps	274	435	2.59	0.84	-1.16	0.11

I = Survey item number as shown in Table 1.
Count = The number of respondents answering a survey item.
Score = The raw score calculated from all survey responses to a single item using the scale: 1 = very easy; 2 = easy; 3 = difficult; and 4 = very difficult.
Raw average = The arithmetic mean of the raw scores.
SD = The standard deviation of the raw scores.
Measure = Item measure in logits.
Error = The standard error of the item measure in logits.

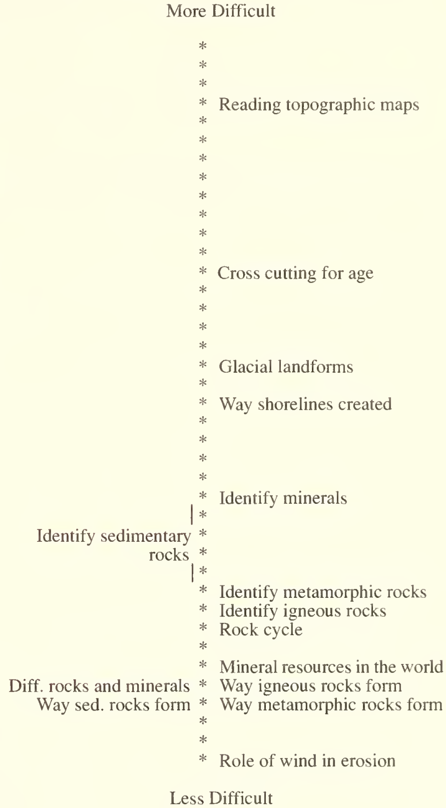
three-dimensional images. The rock cycle could have a low rating because the great length of time involved makes it difficult to visualize when compared to the human life span. The rock cycle is also hard to investigate in a laboratory setting.

Lectures and labs involving the way each rock type (i.e., igneous, sedimentary, and metamorphic) formed generated the same level of interest. Lectures and labs on sedimentary rock formation were rated lower than the labs and lectures involving igneous and metamorphic rock formation. Perhaps sedimentary rock formation is not as dramatic as igneous and metamorphic rock formation. Igneous rocks can be formed by volcanic activity, and metamorphic rocks can be formed when tectonic plates rub against each other. Sedimentary rock formation occurs when sediments are slowly carried into a region of deposition.

If sedimentary regions where substantial wave action occurs were discussed, the lectures and labs on this rock type might not have ranked so low.

Implications of Topics Rated by Difficulty at the Midwestern University. Figure 2 shows that the mineral identification lab was rated as more difficult than the lab on igneous rock identification. Since the mineral identification lab took place before any of the three rock identification labs, and since the ability to identify minerals helps in subsequent labs, the mineral identification lab *should* be more difficult than the igneous rock identification lab to which it contributes. The topic difficulty survey for lecture and lab also reveals another interesting result (Table 3 and Figure 2). The three lectures on rock formation are all rated at about the same difficulty level, but the labs for the identification of each rock type are rated as more difficult. This result seems predictable, since lectures involve listening, while labs involve an active doing, which is often more difficult.

Figure 2. A plot of the 15 Midwestern university class components as a function of item rating displayed on a linear scale corrected with the stochastic model. The standard error of all items was approximately that shown by the error bars for the topic, identify sedimentary rocks. Those items rated as less difficult are listed at the bottom of the chart, while those rated as more difficult are at the top. The full text of all items is given in Table 1.



The high difficulty ratings of the labs for reading topographic maps, age determination, and glacial landforms may be a function of spatial visualization. In the case of glacial landforms, these features must be visualized and identified from two-dimensional maps and photographs, a skill which requires the ability to visualize two-dimensional data in three dimensions. Although shoreline creation is rated as less difficult than these three topics, it is also viewed as more difficult than the remaining eleven topics. Visualization is crucial for understanding the growth and deterioration of shorelines, because deposition along a shore can be a function of water depth. All of the surveyed topics call for some spatial visualization, but the skill level required for the more difficult topics may be higher than that required for other topics.

Rock formation, the differences between rocks and minerals, and the role of wind in erosion were considered far less difficult to understand than the other topics. These topics are simply definitive, while the other topics required the application of definitions and were, therefore, more difficult. At one time, each

Figure 3. A plot of the 15 Southern university class components as a function of item rating displayed on a linear scale corrected with the stochastic model. The standard error of all items was approximately that shown by the error bars for the topic, tectonic plates. Those items rated as more interesting are listed at the bottom of the chart, while those rated as less interesting are at the top. The error bars displayed in Figures 1 and 2 are smaller than those in Figures 3 and 4, because a smaller sample of students was available at the Southern university.

```
                     Less Interesting

 Id. metamorphic rocks * Identify sedimentary rocks
 Identify igneous rocks * Identify minerals
                        *
                        * Reading topographic maps
                        * Way metamorphic rocks form
 Diff. rocks and minerals * Way sedimentary rocks form
                        * Rock cycle
                        * Way igneous rocks form
                        *
                        *
                        * Cross cutting for age
                        *
                        *
                        *
                        * Glacial landforms
                        *
                        *
                        * Groundwater
                        *
                        *
                        *
                        *
                        *
                        *
                        *
                        * |
                        *
                        * Tectonic plates
                        *
                        * |
                        *
                        *
                        *
                        *
                        *
                        *
                        *
                        * Earthquakes
                        *

                     More Interesting
```

type of rock formation might have been viewed as difficult, but by the end of the course, these topics could have been considered relatively easy in comparison to other topics.

A qualitative comparison of Figures 1 and 2 provides added information with regard to the students' attitudes. Although glacial landforms and shoreline creation were viewed as quite interesting, these topics were not listed among the easiest topics by the students. Past experience might have helped the students maintain their interest in these difficult topics. Another explanation might be that a particular presentation technique made a difficult topic more interesting (e.g., very good slides or actively involving the students in the lecture).

The two most difficult topics were also of low interest (i.e., topographic maps and age determination). Predictably, some highly difficult topics might be viewed as not very interesting. Follow-up interviews might reveal if the difficulty of these two topics caused students to view them as uninteresting or if the students lacked interest for other reasons. In the future, students could be randomly selected and interviewed at the end of the course.

Topic Difficulty at Both Institutions. The laboratory identification of rocks and minerals was rated as relatively difficult in comparison to the other topics at both schools (Figures 1 and 3). Furthermore, the ratings for the three rock identification labs at both schools are all clustered together with no clear statistical difference between the sets of labs. Topographic map reading was considered relatively difficult at both institutions. One topic that was rated very differently at the two schools was age determination. Cross cutting relationships were rated as easy at the Southern university (Tables 5 and 6), and this item was also rated fairly high in terms of interest. These results are in marked contrast to what was found at the Midwestern university (i.e., a high difficulty and low interest rating for age determination). A review of course syllabi and discussions with the instructors suggested that the difference was probably related to the complexity of the

Table 4. The list of topics surveyed at a Southern university. Topics 11-15 were labs. One survey asked students to indicate their level of interest for each of the topics by choosing from four possible responses: 1) very disinteresting; 2) disinteresting; 3) interesting; and 4) very interesting. A second survey asked students to indicate how difficult each topic was to understand by selecting from four possible responses: 1) very difficult; 2) difficult; 3) easy; and 4) very easy. The topics presented to students at the second institution are provided in Figures 3 and 4.

1	The difference between rocks and minerals
2	The rock cycle
3	Using cross-cutting relationships to determine the relative age of rocks
4	The way sedimentary rocks are formed
5	The way igneous rocks are formed
6	The way metamorphic rocks are formed
7	Earthquake distribution and occurrence
8	Tectonic plate movement and location
9	Groundwater issues
10	Glacial landforms
11	Identifying minerals
12	Identifying igneous rocks
13	Identifying sedimentary rocks
14	Identifying metamorphic rocks
15	Reading topographic maps

cross-cutting problems posed to students. Simpler cross-cutting scenarios requiring less spatial visualization were used at the Southern university.

The different ratings for the rock cycle is probably related to the presentations used in the two universities. The Midwestern students, who rated the topic as being more difficult (in comparison to the other rated topics), spent a greater amount of lecture time studying the rock cycle. As a result, material of greater complexity was covered. At the second institution, a less-detailed look at the rock cycle was presented. Although differences exist in the relative rating of the rock cycle to other topics at the two institutions, the students rated the rock cycle as easy at both schools.

Interest. The interest data collected from students at both institutions reveals some similar trends across institutions as well as some differences (Figures 2 and 4). At both schools, the labs requiring the identification of rocks and minerals were clumped together based on their interest ratings. The average ratings for these labs at the Midwestern university indicated interest in the labs (Table 2), but the ratings from the Southern university suggested that the students were not interested in these labs (Table 5). A review of the syllabi and topic development in each course suggests that these labs were rated higher at the Midwestern university as a result of over 20 years of course development by the instructor, while the labs at the second university had not been used for nearly as long.

Table 5. The surveyed topics ordered based on the interest ratings of the Southern university students. Items at the top of the table were viewed as more interesting than those listed at the bottom. The sample free statistics calculated for the survey indicate a reliability of 0.90. The raw averages were corrected for non-linearity, and each item's measurement error was calculated so that Figure 3 could be constructed and parametric tests utilized for analysis.

I	Topic	Count	Score	Raw Average	SD	Measure (Logits)	Error (Logits)
7	Earthquakes	52	90	1.75	0.67	2.82	0.28
8	Tectonic plates	52	104	2.07	0.76	1.75	0.27
9	Groundwater	52	118	2.34	0.86	0.79	0.25
10	Glacial landforms	51	120	2.36	0.78	0.52	0.25
3	Cross cutting for age	52	130	2.50	0.76	0.06	0.24
5	Way igneous rocks form	51	132	2.60	0.74	-0.16	0.24
2	Rock cycle	51	134	2.67	0.79	-0.32	0.24
1	Difference between rocks and minerals	52	137	2.68	0.72	-0.35	0.24
4	Way sedimentary rocks form	52	137	2.64	0.70	-0.35	0.24
6	Way metamorphic rocks form	52	139	2.68	0.72	-0.47	0.24
15	Reading topographic maps	52	142	2.75	0.98	-0.64	0.24
11	Identify minerals	52	146	2.82	0.90	-0.87	0.24
12	Identify igneous rocks	52	147	2.84	0.89	-0.93	0.24
13	Identify sedimentary rocks	52	147	2.84	0.89	-0.93	0.24
14	Identify metamorphic rocks	52	147	2.84	0.87	-0.93	0.24

I = Survey item number as used in Table 4.
Count = The number of respondents answering a survey item.
Score = The raw score calculated from all survey responses to a single item using the scale: 1 = very interesting; 2 = interesting; 3 = disinteresting; and 4 = very disinteresting.
Raw average = The arithmetic mean of the raw scores.
SD = The standard deviation of the raw score.
Measure = Item measure in logits.
Error = The standard error of the item measure in logits.

Student interest at both schools was similar in other ways (Figures 1 and 3). For example, topics that the students had some experience with (e.g., running along a shoreline) and/or that were timely (e.g., a major earthquake) were often highly rated. At the Southern university, the sessions on earthquakes, tectonic plates, groundwater, and glaciers were all highly rated, but topographic maps were less interesting.

The topic of groundwater caused some unexpected idiosyncratic responses from the Southern university. Although the students, on average, rated this topic as *interesting,* when all of their responses were reviewed, a number of unexpected *very interesting* ratings were found. This result might indicate a situation that is the opposite of what was observed for topographic maps. At the South-

Table 6. The surveyed topics ordered based on the difficulty ratings of the Southern university students. Items at the top of the table were viewed as less difficult than those listed at the bottom. The sample free statistics calculated for the survey indicate a reliability of 0.93. The raw averages were corrected for non-linearity, and each item's measurement error was calculated so that Figure 4 could be constructed and parametric tests utilized for analysis.

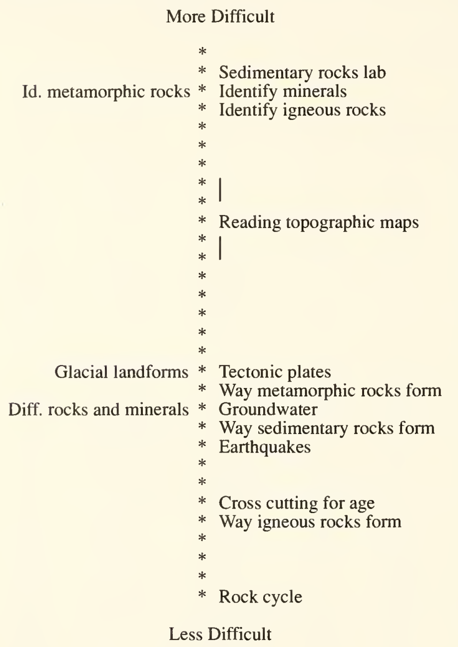
I	Topic	Count	Score	Raw Average	SD	Measure (Logits)	Error (Logits)
2	Rock cycle	55	103	1.87	0.55	1.48	0.28
5	Way igneous rocks form	55	109	1.98	056	1.06	0.27
3	Cross cutting for age	55	110	2.00	0.58	0.93	0.27
7	Earthquakes	56	116	2.07	0.66	0.65	0.26
4	Way sedimentary rocks form	56	118	2.11	0.53	0.51	0.26
1	Difference between rocks and minerals	54	116	2.15	0.63	0.38	0.26
9	Groundwater	56	120	2.14	0.55	0.38	0.25
6	Way metamorphic rocks form	56	121	2.16	0.60	0.31	0.25
10	Glacial landforms	55	120	2.18	0.58	0.23	0.25
8	Tectonic plates	56	123	2.20	0.59	0.18	0.25
15	Reading topographic maps	56	137	2.45	0.87	-0.63	0.23
12	Identify igneous rocks	54	145	2.69	0.84	-1.24	0.22
14	Identify metamorphic rocks	55	150	2.73	0.83	-1.39	0.22
11	Identify minerals	55	150	2.73	0.76	-1.39	0.22
13	Identify sedimentary rocks	55	151	2.75	0.84	-1.44	0.22

I = Survey item number as used in Table 4.
Count = The number of respondents answering a survey item.
Score = The raw score calculated from all survey responses to a single item using the scale: 1 = very easy; 2 = easy; 3 = difficult; and 4 = very difficult.
Raw average = The arithmetic mean of the raw scores.
SD = The standard deviation of the raw score.
Measure = Item measure in logits.
Error = The standard error of the item measure in logits.

ern university, the topic of groundwater was more relevant to the students' real-life experiences than other topics. The instructor said that this particular topic was one which stressed the impact of pollution on groundwater. Also, the issue of groundwater had been a hot political topic in their State.

A comparison of the ratings at the two institutions shows some important similarities as well as some differences. The structure of the class, the experience of the instructor, and the training of the graduate assistants who lead lab sections are all factors that help shape a student's reaction to geology. Although these factors will always remain influential, some particular trends in student attitude were observed across these two schools. Labs generated relatively little interest, but topics that are in the news or that have been experienced by the students are of greater interest than subjects that are less tangible. Finally, at both

Figure 4. A plot of the 15 Southern university class components as a function of item rating displayed on a linear scale corrected with the stochastic model. The standard error of all items was approximately that shown by the error bars for the topic, reading topographic maps. Those items rated as less difficult are listed at the bottom of the chart, while those rated as more difficult are at the top.



schools, student interest in labs and lectures did not differ as a function of rock type taught (i.e., lessons involving igneous, sedimentary, and metamorphic rock).

CONCLUSION

Many plausible reasons exist which might explain the ratings discussed in this paper. Some of these factors are text presentation, the teacher, student attendance, and recent lecture topics. Despite all of these uncontrollable variables, a survey of students' views toward earth science topics reveals some interesting trends and patterns which may be helpful to earth science teachers at all levels and which may provide information for researchers exploring attitudes toward science. Two keys to understanding the ratings appear to be: 1) familiarity with a topic; and 2) the anxiety arising from using spatial visualization for the first time. The topics of greatest interest seem to be those that are in the news, timely, and/or dramatic.

Although the instructors and topics varied, some general trends were observed across institutions. The relatively low interest ratings of the rock identification labs at both schools suggests that the organization of the labs might be changed. Hopefully, doing a lab should be more interesting than passively listening to a lecture. The highly rated topics at both schools seem to be those with which the students can relate. Instructors should try to relate topics to their students' lives and experiences. By doing so, the students' interest may increase. This suggestion might seem to be common sense, but it must be practiced in college science classes. Certainly, a constructivist philosophy of teaching would support this goal (Fraser, 1994).

The results of this study demonstrate that detailed attitudinal data regarding geology topics can be collected, evaluated with a probabilistic model, and used to help instructors understand student perceptions. Whatever the earth science course, this type of analysis can be used to refine courses, ultimately increasing the students' interest and understanding as well as expanding the data base of research on earth science education.

A review of course syllabi, texts, labs, and teaching techniques helps explain how course structure and teaching strategies may have affected topic ratings.

Both instructors lectured to large groups of students, and only a small amount of time was devoted to question and answer. Both instructors used overhead transparencies, board work, and slides.

For these instructors, topics that were tied to a current issue, connected to students' interest, and/or presented in a wide variety of ways were the topics rated as the most interesting. The use of slides, which build upon textbook photos, and simple diagrams seem to decrease difficulty and to increase interest. The instructor at the Southern university felt that his explanation of igneous rock formation was successful because he related the topic to ice cubes melting and forming. He also talked at great length about crystals, such as sugar, dissolving in water when discussing melting.

The instructors and researcher felt that low-rated topics resulted from a lack of analogies, a lack of suitable tie-ins to real life, and, perhaps, to excessive detail when a topic in their particular area of expertise was covered. The instructor at the Southern school felt that the students' view of metamorphic rock formation being more difficult to understand than the formation of other rock types resulted from his use of phase diagrams while discussing the formation of metamorphic rocks. The phase diagrams confused the class. The same instructor also spent a great deal of time talking about stable and unstable minerals. Because no suitable examples were available, the students also found this material confusing.

Another issue was the way the instructors thought students should apply lecture and textbook information in lab. For example, a wide range of techniques that could be used to identify rocks and minerals were presented in lecture and lab. Even though the necessary supplies and reagents were provided to the students, the undergraduates completed the rock and mineral labs by simply memorizing the appearance of the samples.

The collection of attitudinal data from earth science students can provide far-ranging information that is of use to course instructors. The attitudes of the students are easy to evaluate and consider. Since little research has been carried out on earth science classes, almost any study will provide new, publishable information for scientists and science educators. The results of this analysis show that some earth science topics are more interesting to non-majors than others. Since an appreciation (and understanding) of science is so important for all students, this type of study can provide very important information that can be used to improve lab-based science courses.

LITERATURE CITED

Abdullah, W.M.R.B. 1989. One is sufficient or two are necessary: A substitute for "clarifying variables of teaching through factor analysis of Rasch standardized residuals." ERIC Doc. Reprod. Serv. ED 306 312, 28 pp.

Andersen, E.B. 1973. Conditional inference and models for multiple choice questionnaires. Brit. J. Math. Stat. Psychol. 26: 31-44.

__________. 1977. Sufficient statistics and latent trait models. Psychometrika 42: 69-81.

Andrich, D. 1982. An extension of the Rasch model for rating providing both location and dispersion parameters. Psychometrika 47: 105-113.

Armstrong, J.M. 1980. Achievement and participation of women in mathematics: An overview. ERIC Document Reproduction Service ED 184 878, 41pp.

Barndorff-Nielsen, O. 1978. Information and exponential families in statistical theory. John Wiley and Sons, New York, 238 pp.

Bart, H.A. 1991. A hands on approach to understanding topographic maps and their construction. J. Geol. Educ. 39(4): 303-305.

Bezzi, A. 1991. A Macintosh program for improving three dimensional thinking. J. Geol. Educ. 39: 284-288.

Birdd, D.L. 1990. Color me metamorphic. Sci. Teacher. 57(4): 20-25.

Buchwald, C.E. and R.W. Bybee. 1990. Geologists and the reform of science education. J. Geol. Educ. 38: 216-218.

Carpenter, J.R. 1982. A comparison of student attitudes, values, and interests. J. College Sci. Teaching 12: 441-445.

Crawley, F.E., III, and A.S. Coe. 1990. Determinants of middle school students' intention to enroll in a high school science course: An application of the theory of reasoned action. J. Res. Sci. Teaching 27(5): 461-476.

Davis, B., C. Lang, P. Taylor, T. Knight, and B. Sims. 1992. Earth day changes in attitude. Sci. Scope 92: 30-34.

Duncan, O.D. 1984. Notes on social measurement. Russell Sage Foundation, New York, 256 pp.

Eves, R.L. and L.E. Davis. 1987. The "key" method of identifying igneous and metamorphic rocks in introductory laboratory. J. Geol. Educ. 35(3): 150-151.

Finley, F., J. Stewart, and W. Yarroch. 1982. Teachers' perceptions of important and difficult science content. Sci. Educ. 66(4): 531-538.

Fortner, R.W. and V.J. Mayer. 1991. Repeated measures of students' marine and Great Lakes awareness. J. Environ. Educ. 23(1): 30-35.

Fraser, B.J. 1994. Research on classroom and school climate. *In:* D.L. Gabel (Ed.), *Handbook of Research on Science Teaching and Learning,* pp. 493-541, Macmillan, New York, 598 pp.

Happs, J.C. 1982. Glaciers. Sci. Educ. Res. Unit Working Paper 203, 34 pp.

Kelley, P.R. 1979. Combined common person and common item equating of medical science examinations. Paper presented at the Annu. Meeting Amer. Educ. Res. Assoc., San Francisco, California.

Kern, E.L. and J.R. Carpenter. 1984. Enhancement of student values, interests and attitudes in earth science through a field-oriented approach. J. Geol. Educ. 32: 299-305.

Koballa, T., Jr. 1984. Designing a Likert-type scale to assess attitude toward energy conservation: A nine step process. J. Res. Sci. Teaching 20(7): 709-723.

Lombard, A.S. 1982. Effects of reasoning workshops on the teaching strategies of secondary science teachers. Sci. Educ. 66(4): 41-53.

Lucas, K. and J. Dooley. 1982. Student teachers' attitudes toward science and science teaching. J. Res. Sci. Teaching 19(9): 805-809.

Lutgens, F.K. and E.J. Tarbuck. 1989. Essentials of geology, 3rd ed. Merrill, Columbus, Ohio, 378 pp.

McLure, J.W. 1991. The weathering of rocks: Three activities. Sci. Activities 28(1): 23-26.

Meyer, J. 1989. The logic of science — for ninth graders? Stud. Educ. Eval. 15(2): 193-205.

Miall, A.D. 1982. Recent developments in facies models for siliciclastic sediments. J. Geol. Educ. 30(4): 222-240.

Monk, J.S. 1984. An examination of methods used to generate daily group scores from single-item-per-subject data collected in intensive time-series designs. J. Res. Sci. Teaching 21(3): 315-324.

Newman, W.L. 1983. Geologic time. U.S. Geological Survey. ERIC Doc. Reprod. Serv. ED 241 335, 22 pp.

Oringer, R.H. 1985. Continental jigsaw. Sci. Activities 22(2): 39-40.

Pasteris, J.D. 1983. Value of reflected light microscopy in teaching. J. Geol. Educ 31(1): 17-22.

Perdue, P.K. 1991. Science is an action word! Grades 1-3. Scott, Foresman and Company, Glenview, Illinois, 84 pp.

Preece, P.F.W. 1979. Objective measurement in education: The Rasch model. School Sci. Rev. 60(213): 770-773.

Raatz, U. 1985. Investigating dimensionality of language tests — A new solution to an old problem. *In: Practice and Problems in Language Testing 8,* pp. 123-135, ERIC Doc. Reprod. Serv. ED 266 638.

Rasch, G. 1980. Probabilistic models for some intelligence and attainment tests. Univ. Chicago Press, Chicago, Illinois, 199 pp.

Reckase, M.D. 1979. Unifactor latent trait models applied to multifactor tests: Results and implications. J. Educ. Statist. 4(3): 207-230.

Renner, J., M. Abraham, and H. Birnie. 1985. Secondary school students' beliefs about the physics laboratory. Sci. Educ. 69(5): 649-663.

Rost, J. 1988. Measuring attitudes with a threshold model drawing on a traditional scaling concept. Appl. Psychol. Measurement 12(4): 397-409.

Sherman, R. and K. Stone. 1992. Resources for teaching about energy in the social studies classroom. Soc. Educ. 56(1): 27-30.

Smith, W.S. and C.I. Litman. 1979. Early adolescent girls' and boys' learning of a spatial visualization skill. Sci. Educ. 63: 671-676.

Tarbuck, E.J. and F.K. Lutgens. 1993. The earth: An introduction to physical geology, 4th ed. Macmillan, New York, 654 pp.

Thorndike, E.L. 1904. An introduction to the theory of mental and social measurements. Teachers Coll., Columbia Univ., New York, 212 pp.

Thurstone, L.L. 1925. A method of scaling psychological and educational tests. J. Educ. Psychol. 17: 446-457.

__________. 1926. The scoring of individual performance. J. Educ. Psychol. 17: 446-457.

__________. 1928. Attitudes can be measured. Amer. J. Sociol. 33: 529-554.

Towse, P.J. 1983. Do new science courses improve attitudes toward science? A study in Lesotho. Sci. Educ. 67(2): 159-169.

Wongbundhit, Y. 1985. Item banking procedure and quality control in Dade County Public Schools. Paper presented at the Annu. Meeting Amer. Educ. Res. Assoc., Chicago, Illinois.

Wright, B.D. 1991. Factor item analysis versus Rasch item analysis. Rasch Measurement 5(1): 134-135.

__________ and J.M. Linacre. 1991. Bigsteps computer program. MESA Press, Chicago, Illinois, 80 pp.

__________ and G.N. Masters. 1982. Rating scale analysis. MESA Press, Chicago, Illinois, 206 pp.