JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 33, NO. 4, PP. 369–391 (1996)

# Spatial Abilities of High-School Students in the Perception of Geologic Structures

Yael Kali and Nir Orion

*Department of Science Teaching, Weizmann Institute of Science, Rehovot 76100, Israel*

## Abstract

The specific spatial abilities required for the study of basic structural geology were characterized by quantitative and qualitative data analysis. A geologic spatial ability test (GeoSAT) was developed and administered to 115 comprehensive high-school students. Six of these students were interviewed. An analysis of students' incorrect answers revealed two types of answers: (a) nonpenetrative answers, which were based on external exposures of the structure; and (b) penetrative answers, which indicated attempts at representing internal properties of the structure. Students who tended to give penetrative incorrect answers performed significantly higher than students who tended to give nonpenetrative incorrect answers. The reasoning of students for these types of answers, as determined by interviews, supported the initial assumption that these answers were given by students with different levels of ability mentally to penetrate the image of a structure, which was named *visual penetration ability* (VPA). The interview findings indicated that the VPA is one of two complementary factors needed to solve the problems of GeoSAT; the other factor is the ability to perceive the spatial configuration of the structure. It is concluded that the teaching and learning process should provide students with assistance in both of these areas.

A great amount of research has been dedicated to the investigation of spatial abilities. One of the main reasons for this emphasis is the importance of these abilities in fields such as natural sciences, geometry, engineering, and architecture (McGee, 1979). In the earth sciences, spatial abilities play a fundamental role in several different topics. This role was considerably emphasized in Chadwick's (1978) description of geologic thinking. Chadwick, a geologist and psychologist, claimed that:

> For efficient and geologically adaptable thinking, one prerequisite is probably of universal value, whatever the nature of the geological content. This is the skill for thinking in three dimensions, for visualizing shapes in the mind's eye, rotating, translating and shearing them, and for imagining complex changes over time in the form of a cinematographic visual image. (p. 144)

One of the areas of earth sciences that requires spatial abilities in particular is structural geology. This branch of the earth sciences deals with deformational structures of the earth's crust and their relation to internal forces of the earth. To study these relations, the ability to

Published by John Wiley & Sons, Inc. CCC 0022-4308/96/040369-23

perceive the spatial configuration of various structures and envision the shapes of their different cross-sections is required.

The development of children's conception of sectional properties of objects was studied by Piaget and Inhelder (1956). In their investigation, 4- to 12-year-old children were presented with plasticine models of geometric objects and asked to predict the shapes of different bisections of those objects. Piaget and Inhelder described four characteristic developmental stages, from a first stage where children tend to confuse external and internal parts of the object, to the fourth stage where children are usually able to predict the correct shapes of the bisections.

A clear difference exists between the homogeneous plasticine objects of Piaget and Inhelder and the more complex layered structures involved in geology. Therefore, it is not surprising that even in high schools, students experience considerable difficulties in the perception of geologic structures and in envisioning different cross-sections in these structures. The existence of difficulties induced by spatial ability requirements is a characteristic of various topics related to different disciplines. Some of them are geometry (Battista, Wheatley, & Talsma, 1989; Ben-Chaim, Lappan, & Houang, 1985; Cooper, 1992), chemistry (Dyche, Mclurg, Stepans, & Veath, 1993; Seddon & Moor, 1986; Small & Morton, 1983; Tuckey & Selvaratnam, 1993), biology (Russell-Gebbett, 1984, 1985), astronomy (Broadfoot, 1993), and engineering graphics (Rodriguez, 1990; Wiley, 1990). A majority of the studies that investigated the role of experience and the effect of instruction on spatial skills indicated that these skills can be improved through learning experiences (Baenninger & Newcombe, 1989; Ben-Chaim, Lappan, & Houang, 1988; Kiser, 1990; Lord, 1985, 1987; Smith & Schroeder, 1981).

A greatly debated issue is the existence of a gender difference in spatial abilities. Although considerable dispute surrounds the magnitude of this difference, the age of its arousal, and its sources, many researchers agree that the spatial abilities of males are more highly developed than those of the females (Baenninger & Newcombe, 1989; Hyde, 1990; Johnson & Meade, 1987; Linn & Petersen, 1985; Maccoby & Jacklin, 1974; McGee, 1979). However, evidence exists that instructional programs induce similar improvement rates for males and females (Ben-Chaim et al., 1988).

To develop curriculum materials that will assist female and male students in acquiring spatial skills needed for any specific field, it is necessary to investigate the specific characteristics of the spatial abilities required. Several different general classification systems describing types of spatial abilities have been defined. However, it is generally agreed that spatial abilities include the following characteristic categories:

1. The ability to recognize and comprehend the relationships between the various parts of a configuration and one's own position. This category corresponds the "spatial orientation" of McGee (1979) and the "spatial perception" of Linn and Petersen (1985).
2. The ability to generate an image and operate various mental manipulations on this image. This category corresponds the "spatial visualization" of both McGee (1979) and Linn and Petersen (1985). However, the latter authors suggest a third category, mental rotation, whereas McGee referred to the mental rotation as a specific type of manipulation included in the spatial visualization category.

The purpose of the current study was to characterize the specific spatial abilities required in basic structural geology studies, through the examination of the performance of high-school students in solving structural geology problems. To characterize these specific abilities, the following questions were considered:

1. What are the typical answers students give in solving such problems?
2. What are students' reasoning behind different types of answers?

3. What is the distribution pattern of student performance in solving basic structural geology problems, which require spatial abilities?
4. Can relationships be found between answer type and achievement?

## Method

### *Sample*

The sample consisted of 115 students: 44 males and 70 females (1 student did not indicate his or her gender). This sample was composed of three classes which were randomly selected out of eight 10th-grade (age 16) classes, in a comprehensive Israeli high school. The sample students learned some geologic topics in their geography lessons prior to this study, and did not receive any geology lessons during the whole period of the study. All students of the sample were tested, and 6 of them (3 females and 3 males) were interviewed as well.

### *Instrument*

A geologic spatial ability test (GeoSAT) was developed for the current study. The development process of the test included the design of a pilot version, which consisted of 15 multiple-choice problems. This test was administered to 64 10th-grade students, and enabled gaining preliminary quantitative data of high-school students in solving spatial problems related to geologic structures. To obtain further insight into students' perceptions and difficulties concerning those abilities, it was necessary to design the current version of GeoSAT as an open-ended test including several comparable types of problems.

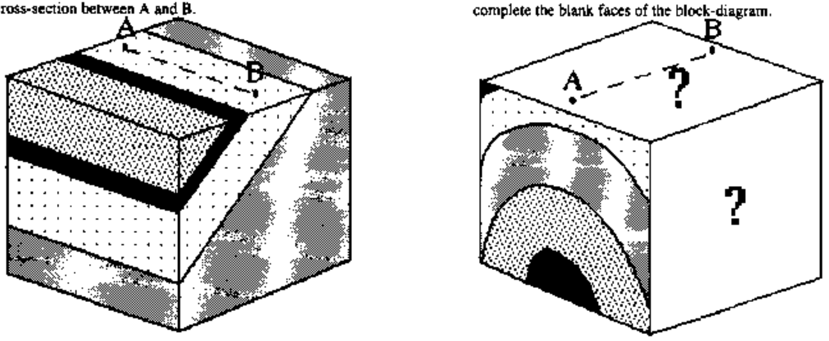
The current version is composed of 13 open-ended problems that require spatial perception of geologic structures. According to the pilot version, this number of problems provides students with reasonable time for encountering the whole test in a period of one lesson. The test includes three types of problems which represent different types of spatial tasks needed in structural geology and, according to the pilot version, should reveal students' difficulties. These problems are grouped in the following subtests (Figure 1):

1. Cross-section subtest, including four problems which require drawing cross-sections of structures presented as block diagrams (Figure 1a).
2. Completion subtest, including four problems which require completing block diagrams that reveal only a single face (Figure 1b).
3. Construction subtest, including five problems in which two cross-sections and their location on a very simplified geologic map are given. The students are required to draw a third cross-section at a specified location on the map (Figure 1c).

Each subtest is based on the same geologic structures, which include inclined flat layers, two types of horizontal folds (upright synclinal and anticlinal), and a plunging anticlinal fold (Figure 2).

The test includes an instruction sheet, which gives an illustrated explanation of the concept *cross-section*, designed for non–earth sciences students. In addition, the following guidelines are listed:

1. The problems might have more than one correct answer.
2. The layers are continuous and have consistent thicknesses.
3. The block diagrams can be regarded as cut out of larger three-dimensional structures.



1.a. Cross-section problem - students were required to draw the vertical cross-section between A and B.

1.b. Completion problem - students were required to complete the blank faces of the block-diagram.

1.c. Construction problem - students were required to draw the cross-section between D and C, beeing given the following data.

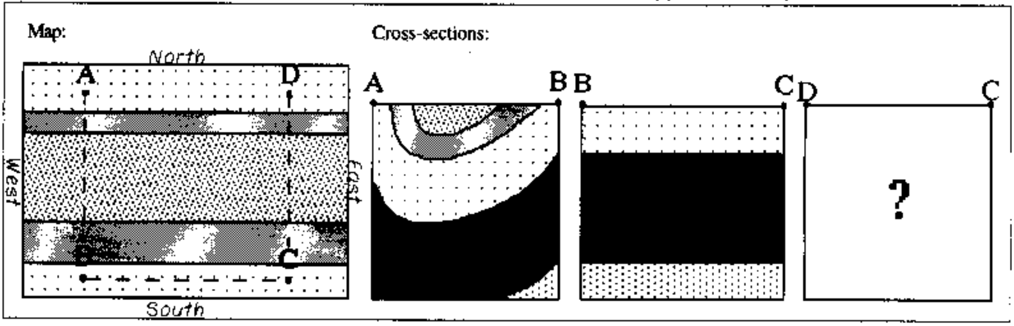


*Figure 1.* Examples of the three types of problems included in GeoSAT.

## *Validation and Characterization of the Instrument.*

*Statistical Notice.* Since nonnormal score distributions were involved, nonparametric procedures were used for characterization of the instrument. Therefore, Spearman's rank was used for correlation and Friedman's two-way analysis of variance by ranks was used for comparing *k*-related samples.

1. Cronbach's α reliability coefficient for the whole test was .90, with a range from .71 to .84 in the three subtests.
2. Spearman's rank correlation coefficients (*Rho*) indicated that a significant positive correlation exists between performance on each of the subtests (Table 1).
3. Friedman's two-way analysis of variance for related samples (the three subtests comprise one direction, and the students who completed these subtests comprise the other) indicated that significant differences exist between performance on the subtests ($df = 2$, $\chi_r^2 = 7.55$, $p < .05$).

   A similar procedure indicated differences between items (structures) within the cross-section subtest ($df = 3$, $\chi_r^2 = 52.34$, $p < .001$) and within the construction subtest ($df = 4$, $\chi_r^2 = 9.87$, $p < .01$).

   Friedman's multiple comparisons between the subtests and between the items enabled definition of the following hierarchic ranking order of difficulty levels (from the easiest to most difficult): (a) cross-section and completion problems based on simple structures, including the two kinds of horizontal fold structures, and the inclined flat layers structure; (b) cross-section problems based on complicated structures,

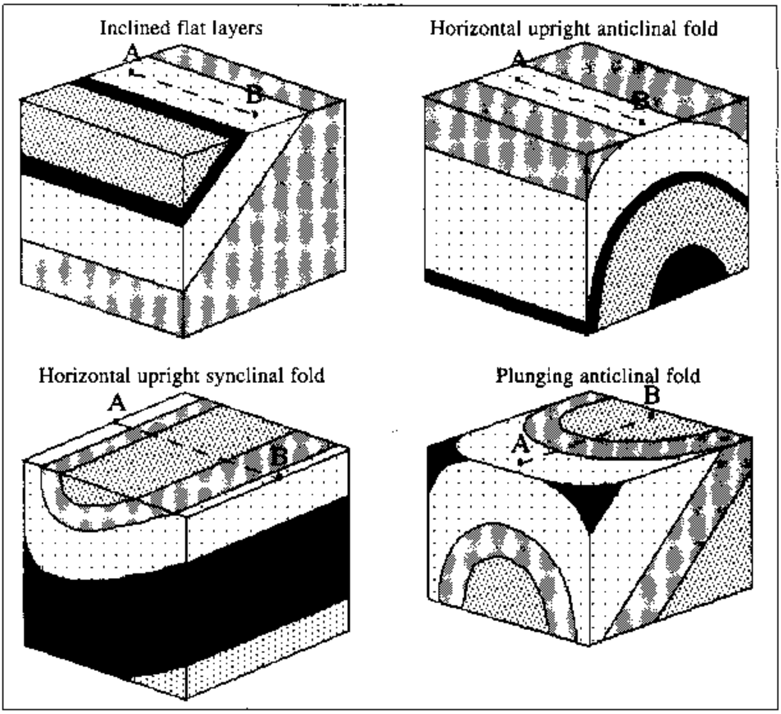


*Figure 2.* Types of structures included in each subtest.

including the plunging fold structure; (c) construction problems based on simple structures; and (d) construction problems based on complicated structures.

4. Expert judgment indicated that the three item types of GeoSAT belong to the *spatial visualization* category described by Linn and Petersen (1985) and McGee (1979). Further, according to the classification of figural tests (Eliot, 1980), GeoSAT's problems should be considered as manipulative spatial tasks rather than matching tasks.

Table 1
*Correlation between the Three Subtests*

| | Cross-section | Completion | Construction |
|---|---|---|---|
| Cross-section | 1 | | |
| Completion | *Rho* = .564<br>*df* = 80, (*n* = 82)<br>*t* = 6.727** | 1 | |
| Construction | *Rho* = .542<br>*df* = 32, (*n* = 34)<br>*t* = 3.648** | *Rho* = .488<br>*df* = 32, (*n* = 34)<br>*t* = 3.163* | 1 |

*Note*. The reduction in the sample size is caused by the score assessment method (see data analysis).
*$p < .01$ **$p < .001$.

### *Procedure*

Data were collected through testing and interviewing the sample students.

*Testing.* In each of the classes the students were initially given a 5-min period during which the instructor read the instructions out loud. This was followed by a 40-min period in which students answered the problems of the test. After this period, the forms were collected from all of the students, including those who did not complete the entire test.

*Interviewing.* The interviews were conducted 3 months after the test was given. No geology lessons or treatments for enhancing spatial skills were given to the students during this period. The interviews were based on individual reexamination of the four cross-section problems of GeoSAT, which the students had answered in the test. The cross-section problem type was chosen for the interviews following an analysis of student answers in the test, which revealed typical incorrect answers to these problems. The objective of the interviews was to characterize these answers according to the reasoning of students. Therefore, the criterion for choosing students for the interviews was based on their answers in the test: Students who tended to give a specific type of incorrect answer to most of the problems in the test were chosen for the interviews. The interviews were conducted as conversations which lasted 20 to 60 min with each student, depending on his or her level of cooperation. This interaction included two stages. In the first stage, students were given an opportunity to reexamine their former answers to the four problems, to indicate whether they still stood behind them, and to revise answers which they thought were mistaken. At the second stage, students were asked again to reconsider their answers; this time, they were given a multiple-choice form of those problems. At this stage, they were asked to confront their multiple-choice answer, with their revised answer of the first stage. When the multiple-choice answers did not agree with the revised answers, students were asked to explain their reasons for giving each of the answers, and the reasons for changing their minds in the second stage. The multiple-choice form was designed especially for the interviews, and included the four cross-section problems of GeoSAT, each problem with four choice possibilities. The choices included 1 correct answer and 3 distracters, which represent typical incorrect answers that were found in the analysis of students' answers in the test (an example is given in Students' Reasoning) (Figure 12). The interviews were tape recorded, transcribed, and then analyzed.

### *Data Analysis*

*Score Assessment.* To evaluate students' achievements on different problem types, it was necessary to assess scores for each subtest and for the entire test. Since many tests were not entirely completed, scores were assessed for each subtest through the percentage of correct answers out of the bulk of completed answers, providing that at least three problems were answered. Test forms in which fewer than three problems were completed for a certain subtest were not taken into account in the score analysis of that particular subtest. This method was based on the assumption that uncompleted answers within the bulk of the completed answers were not understood, and that uncompleted answers following this bulk were simply not reached by the students. Therefore, following the bulk of completed answers, no answer was not considered an incorrect answer. Consequently, the sample size was reduced from 115 students tested to 101 students in the cross-section subtest, 82 in the completion subtest, and 34 in the construction subtest. This method might have created a bias if a correlation had existed between the percent of completed answers and performance. However, the correlation between these factors was close to zero, indicating that this is not a source for bias.

*Checking for Correctness.* The cross-section and construction subtests included a single possible correct answer. Answers were considered to be correct if the general pattern of the layers was correct. Inaccuracies in depicting the thickness of layers, or within their location in the cross-section, were ignored. However, in the completion subtest, where an infinite number of possible correct answers existed, correct answers were considered to be those completions that created reasonable three-dimensional structures (Figure 3a).

Answers were considered to be incorrect if they presented two-dimensional completions. Such incorrect answers include completions with discontinuities between layers exposed on the different faces of the block diagram, completions based on face duplication with or without the usage of mirror symmetry, or completions based on continuation of lines without consideration of the block diagram as a three-dimensional object (Figure 3b).

*Analysis of Incorrect Answers.* Students' answers in the cross-section subtest enabled clear identification of different types of incorrect answers. This identification enabled insight into

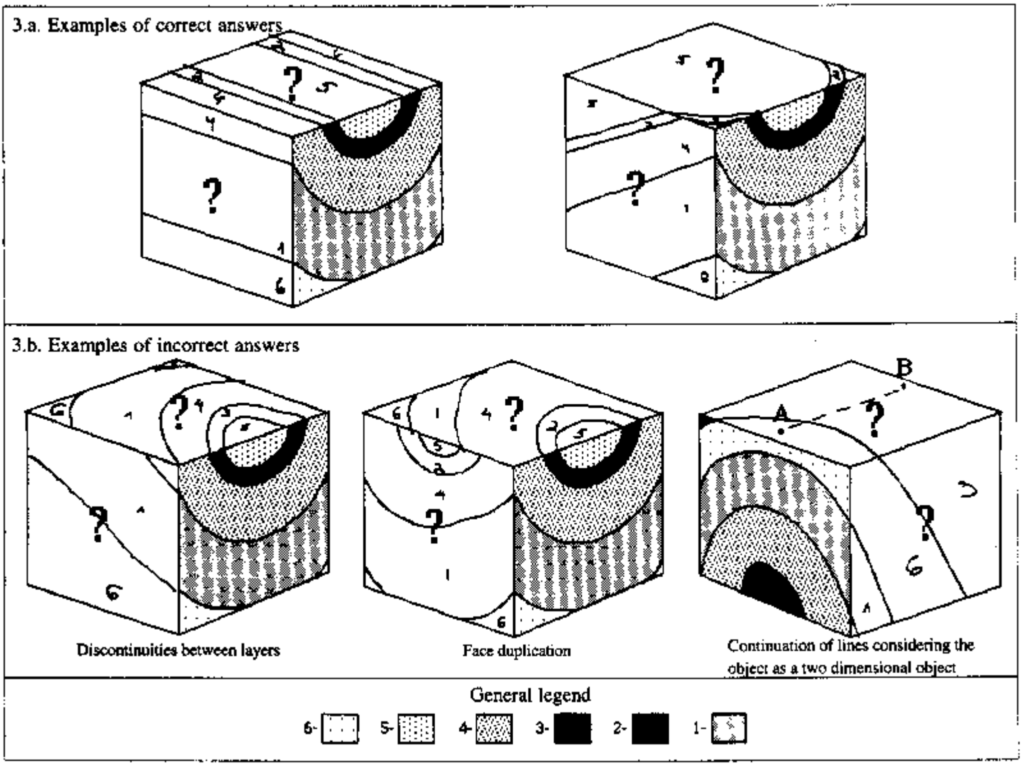


*Figure 3.* Examples of different correct and incorrect answers of completion problems. Students were required to complete the blank faces of the block diagram.

students' thinking processes and provided the means for examination of relationships between the tendency of students to use certain types of incorrect answer, and achievements on the different subtests.

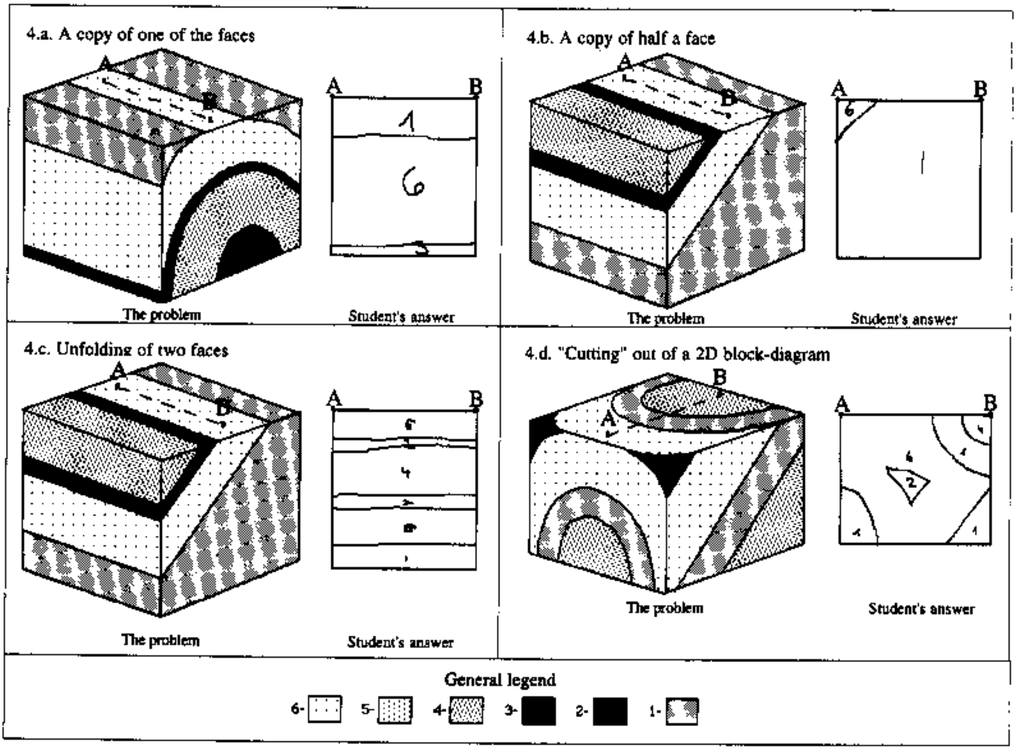
*Statistical Analysis.* Frequency distributions of scores were described using Pearson's probability distribution chart, which relates skewness and kurtosis coefficients with frequency curves (Pearson & Hartley, 1976). Because of the nonnormal distributions that were found, all statistical procedures were restricted to nonparametric techniques. Since the variables were at least of ordinal scale, the following procedures were used (Seigel & Castellan, 1988): Mann–Whitney *U*-test for comparisons between two independent samples and Spearman's rank for correlation.

## Results

### *Types of Incorrect Answer*

Two types of incorrect answer were found in the cross-section subtest:

1. Incorrect answers that are based on external patterns exposed on the visible faces of the block diagram. These answers were based on either one, two, or three of these faces. Different examples of such answers follow (Figure 4): (a) answers that show a copy of



4.a. A copy of one of the faces

4.b. A copy of half a face

4.c. Unfolding of two faces

4.d. "Cutting" out of a 2D block-diagram

*Figure 4.* Examples of incorrect answers based on external patterns. Students were required to draw a vertical cross-section through A and B.

one of the faces of the block diagram (about 75%) (Figure 4a); (b) answers that show a copy of the remaining half of the bisected face (about 2%) (Figure 4b); (c) answers that are based on unfolding of two faces (about 5%) (Figure 4c); and (d) answers that are derived from imagined cutting of the block diagram as a two-dimensional object, and combining segments of the three external faces of the block diagram (about 15%) (Figure 4d).

It was postulated that these types of answers reflect students' difficulties in mental penetration into the structure. Therefore, such answers were termed *nonpenetrative* answers.

2. Incorrect answers that indicated an attempt to present interior properties of the block diagram. Different examples of such incorrect answers follow (Figure 5): (a) answers including vertical continuation of layers, which are exposed at the top of the block diagram, and horizontal continuation of the layers, which are exposed at one of the sides of the block diagram (about 85%) (Figure 5a); and (b) answers including layer continuation from only one of the faces of the block diagram, usually as straight lines (about 10%) (Figure 5b).

It was postulated that these types of answers were given by students who made attempts mentally to penetrate into the structure, and thus deduced the shape of the cross-section. Therefore, these answers were termed *penetrative* answers.

### *Frequency Distribution of Incorrect Answer Types*

Figure 6 shows the distribution of students by the percentage of penetrative answers within the incorrect answers they gave. Only 7 of the 101 students who completed this subtest gave correct answers for all of the four cross-section problems, and are therefore not included in this distribution, reducing the sample to 94 students.

Figure 6 indicates that most of the students (86% of the sample) gave consistent types of incorrect answers: 44 students (47%) gave 100% nonpenetrative incorrect answers, and 37 students (39%) gave 100% penetrative incorrect answers.

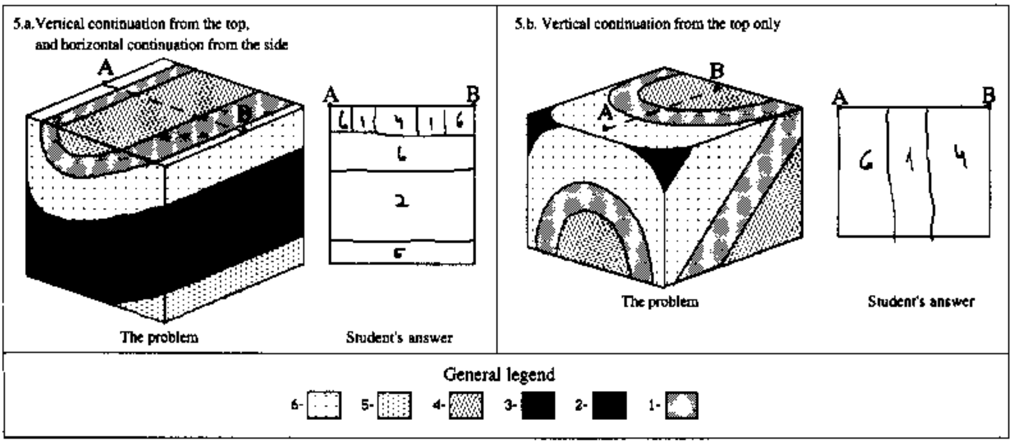


*Figure 5.* Examples of incorrect answers which indicate an attempt to present interior properties of the structure. Students were required to draw a vertical cross-section through A and B.

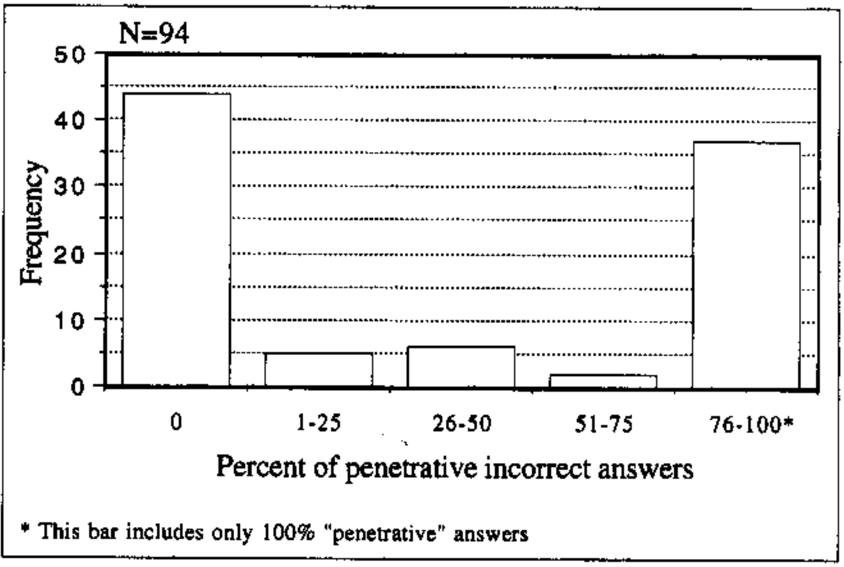


*Figure 6.* Distribution of the percentage of penetrative incorrect answers given by students.

### *Achievement Differences between Types of Students*

Comparisons between the achievements of students who gave 100% nonpenetrative incorrect answers and those who gave 100% penetrative answers in the three subtests were analyzed by Mann–Whitney *U*-test (Table 2). Significant differences favoring students who tended to give penetrative incorrect answers were found in performance on both the cross-section and the completion subtests. This difference suggests that a positive correlation should exist between the percentage of penetrative incorrect answers and performance on these subtests. This correlation was found using Spearman's rank correlation on the cross-section subtest ($Rho = .525$, $df = 92$,

Table 2

*Achievement Differences between Students Who Tended to Give Penetrative Incorrect Answers and Those Who Tended to Give Nonpenetrative Incorrect Answers in the Three Subtests*

| | Non-P[a] | | P[b] | | |
|---|---|---|---|---|---|
| Subtest | *n* | Mean rank of scores | *n* | Mean rank of scores | *z* |
| Cross section | 44 | 30.65 | 37 | 53.31 | 4.319** |
| Completion | 44 | 36.65 | 36 | 45.21 | 1.705* |
| Construction | 15 | 13.20 | 11 | 13.91 | .244 |

*Note.* The samples include only those students who gave consistent incorrect types of answers in the cross-section subtest. The reduction in the sample size through the subtests is caused by the score assessment method (see data analysis). P = penetrative.

[a] Students with 100% nonpenetrative incorrect answers in the cross-section subtest.

[b] Students with 100% penetrative incorrect answers in the cross-section subtest.

$^{*}p < .05$ $^{**}p < .001$.

$t = 5.917, p < .001$), and on the completion subtest ($Rho = .273, df = 73, t = 2.707, p < .01$). The higher achievements of students who tended to give penetrative answers suggests that the ability to use a penetrative strategy is advantageous for performance on this test.

### *Frequency Distributions of Scores*

The frequency distributions of the scores on each subtest and on the entire test, and the distributions of the female and male subsamples are shown in Figure 7. It can be seen that in each of the three subtests and on the entire test, the score distributions of the entire sample show bimodal patterns. According to Pearson's probability distribution chart, these distributions belong to the U-shape category of distributions (Table 3). However, the distributions of the entire sample are actually superpositions of the female and male subsamples. The separation of these subsamples reveals that in the cross-section and the completion subtests, the bimodal pattern characterizes only the females' distribution, while the males' distributions tend toward high scores (negatively skewed) exhibiting Pearson's J-shape pattern (Table 3). In the construction subtest, however, the males' scores are those that exhibit the U-shape pattern, while the females' score distribution tends toward low scores (positively skewed) and also exhibit Pearson's J-shape pattern (Table 3).

### *Gender Differences*

Comparisons between the performance of females and males were analyzed by Mann–Whitney's *U*-test (Table 4). Significant differences favoring males were found in each of the subtests.

## Students' Reasoning

The objectives of the interviews were to obtain further insight into different types of answers which were given in the test, and to characterize these answers by students' reasoning. Six students from the tested sample were chosen for the interviews. The rationale was to select students who tended to give a specific type of incorrect answer to the four cross-section problems of the test. Consequently, 3 students (2 females and 1 male) who tended to give penetrative answers and 3 students (1 female and 2 males) who tended to give nonpenetrative answers were interviewed. The students did not change this tendency during the interview.

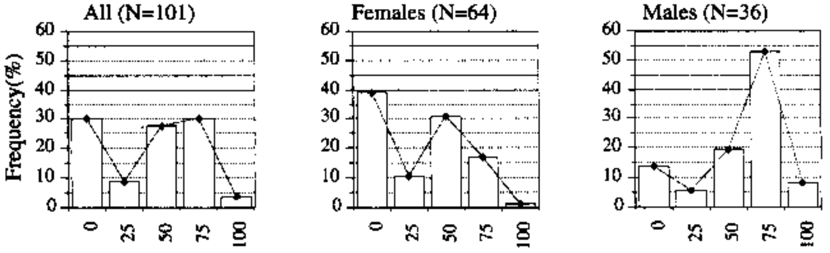
### *Reasoning for Correct Answers Given by Students Who Tended to Give Penetrative Incorrect Answers*

The reasoning of these students for correct answers was usually based on a description of the structure. Their point of view was that if they described the structure's shape, the interviewer would automatically understand why their drawings were appropriate. These students tended to be very confident of their answers. An example can be seen in Chaya's (female) reasoning, in a situation where she was asked to reexamine her former correct answer (Figure 8).

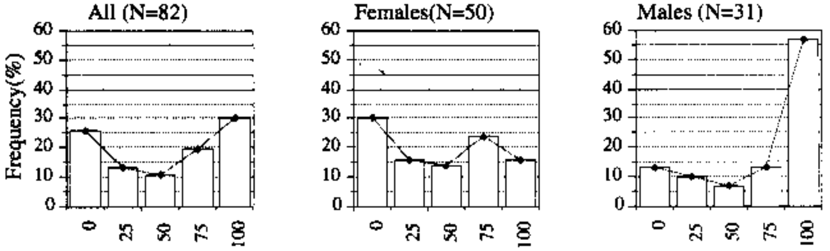
### *Students' Reasoning for Penetrative Incorrect Answers*

Students' reasoning for penetrative incorrect answers usually indicated an inability to perceive the spatial configuration of the structure, and were based based on an analytic ap-

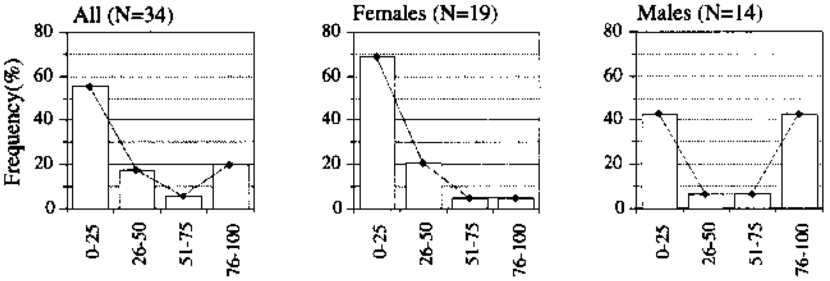
7.a. Score distributions on the cross-section sub-test



7.b. Score distributions on the completion sub-test



7.c. Score distributions on the construction sub-test



7.d. Score distributions on the entire test

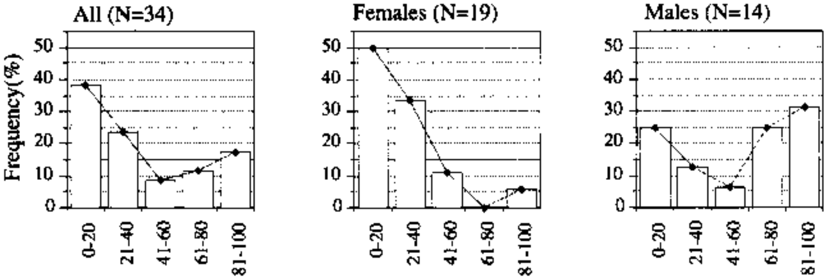


*Figure 7.* Frequency distributions of scores on each subtest, and on the entire test.

Table 3
*Types of Score Distribution on Each of the Subtests and on the Entire Test, Given for the Whole Sample, Females and Males*

| Test | Sample | Skewness coefficient ($b_1$) | Kurtosis coefficient ($b_2$) | Pearson's distribution type |
|---|---|---|---|---|
| Cross-section subtest | All | 0.110 | 1.612 | U-shape |
| | Females | 0.219 | 1.594 | U-shape |
| | Males | 0.701 | 2.611 | J-shape |
| Completion subtest | All | 0.173 | 1.339 | U-shape |
| | Females | 0.420 | 1.614 | U-shape |
| | Males | 0.306 | 2.150 | J-shape |
| Construction subtest | All | 0.675 | 1.952 | U-shape |
| | Females | 0.994 | 3.083 | J-shape |
| | Males | 0.006 | 1.350 | U-shape |
| Entire test | All | 0.553 | 1.743 | U-shape |
| | Females | 1.447 | 4.188 | J-shape |
| | Males | 1.391 | 0.357 | U-shape |

proach, which shows an attempt to deduce the interior parts of the structure. The students who gave penetrative answers tended to view their answer as one of many possible answers, and were somewhat hesitant. An example of this analytic penetrative approach and the hesitant attitude can be seen in another part of the interview with Chaya (Figure 9).

### *Students' Reasoning for Nonpenetrative Incorrect Answers*

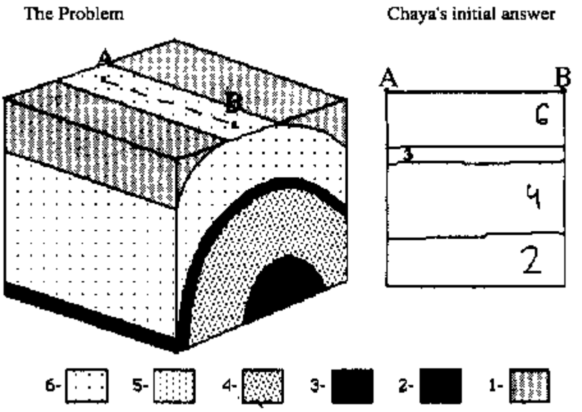
The reasoning of the students was completely based on external patterns exposed on the faces of the block diagram, with an inability to envision internal parts of the structures. Nonpenetrative answers were given by students who were not able to perceive the spatial configuration of the structures shown, as well as by students who were able to perceive these structures. An example of reasoning based on external patterns, given by a student who was not able to perceive the spatial configuration of the structure, can be seen in part of the interview with Yaniv (male) (Figure 10).

An example of a nonpenetrative student who was able to perceive the structure but was not

Table 4
*Comparisons between the Scores of Females and Males*

| | Females | | Males | | |
|---|---|---|---|---|---|
| Subtest | *n* | Mean rank of scores | *n* | Mean rank of scores | *z* |
| Cross section | 64 | 41.27 | 36 | 63.32 | 3.682*** |
| Completion | 50 | 43.38 | 31 | 60.51 | 2.858** |
| Construction | 19 | 13.82 | 14 | 20.38 | 2.031* |

*Note.* The reduction in the sample size is caused by the score assessment method (see data analysis).
$^*p < .05$ $^{**}p < .01$ $^{***}p < .001$.



S: *(after a very short glance)* It's correct.
I: Can you explain why you think this is what the cross-section looks like?
S: It's clear that the layers are straight, because they are straight here as well *(points on the left face)*, but since there is a long hill here in the middle, and the cross-section is at the top of the hill, we'll see layers 4 and 2 as well.

*Figure 8.* An example of students' reasoning for correct answers.

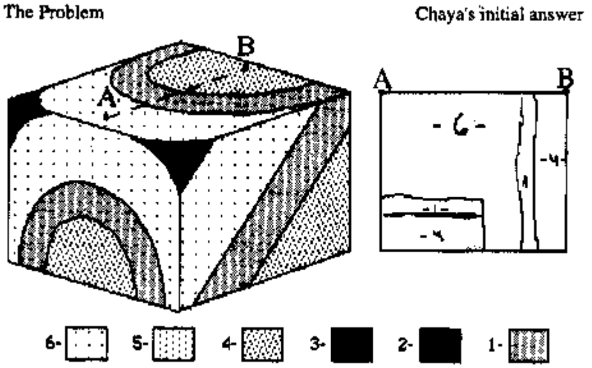
able to envision the cross-section can be seen in part of the interview with Doron (male) (Figure 11).

A phenomenon which was revealed in the interviews was that when students who gave nonpenetrative answers were not able to perceive the spatial configuration of the structure, they had difficulties in identifying the correct cross-section in the multiple-choice form of the problem. However, whenever they were able to perceive the structure, they identified the correct cross-section, although it was different from their initial answer. These students indicated that their initial answers were a result of confusion caused by the external parts of the block diagram. An example of an identification of the correct cross-section in the multiple-choice form of the problem, by a student who initially gave an incorrect nonpenetrative answer, can be seen in another part of the interview with Doron, who was given the same problem shown in Figure 11, this time presented in a multiple-choice form (Figure 12).

Hence, the types of reasoning given by students for the different types of answers agree with the postulations that were made in the classification of these types: students' reasoning for penetrative answers indicates attempts to deduce internal properties of the structures, and the reasoning for nonpenetrative answers shows a dependency of students on the patterns exposed on the external faces of the block diagram. In addition, another factor influencing students' answers is the extent of their ability to perceive the spatial configuration of the structure. Correct answers are produced only when the structure is perceived by students who are able to operate a penetrative strategy.

## Discussion and Conclusions

An analysis of students' incorrect answers indicates that these answers divide the population in an almost dichotomized manner, into students who consistently gave nonpenetrative



I: Can you explain why you think this is what the cross-section looks like?

S: What I tried to do here is... I don't know how deep this goes inside *(points on layer 1 in the upper face)* ... this is what bothers me... so I took these two *(points on the left and upper faces)*. If we look only on the upper face... *(starts drawing vertical lines which continue the layers from the upper face).*

I: Do you think the layers must be vertical?

S: I can't really know *(proceeds in drawing horizontal lines which continue the layers from the left face and reach the vertical layers she drew before, creating the drawing below).*

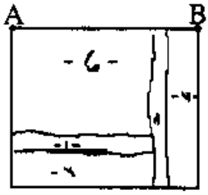
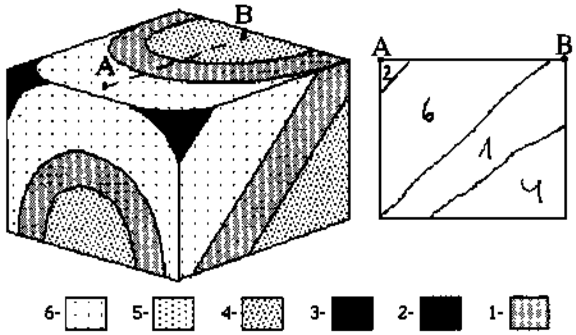


*Figure 9.* An example of students' reasoning for penetrative incorrect answers.

incorrect answers and those who consistently gave penetrative incorrect answers. The higher achievements of the latter students on the cross-section and completion subtests suggest that a certain ability, related to the incorrect penetration type of answers, plays a critical role in solving problems of these tests. We refer to this ability as VPA.

The suggestion that VPA is involved in solving the problems of GeoSAT, and its reflection on different types of incorrect answers, were supported by the interviews: Students' reasoning for penetrative incorrect answers indicated attempts to deduce internal properties of the structures, while the reasoning for nonpenetrative answers showed a dependency of students on the patterns exposed on the external faces of the block diagram, and an inability mentally to penetrate into the structure.

The fact that students with high VPA gave incorrect answers to some of the problems suggests that additional factors which prevented these students from answering correctly must be involved. Analysis of the interviews reveals that such a factor is the ability to perceive the spatial configuration of the layers comprising the structure. The effect of this factor was particularly noticeable in interviews with students who tended to give penetrative answers. These students drew the correct cross-sections easily when they were able to perceive the spatial



The problem

Yaniv's initial answer

I: Can you explain why you think this is what the cross-section looks like?

S: Because the cross-section is from here *(points on the right side of the block-diagram)*, parallel to this side.

I: Can you describe this structure?

S: Yes, *(gives a description of the external appearance of the layers)*

I: Do you think the layers are connected inside the structure?

S: They probably are.

I: Can you describe in which manner?

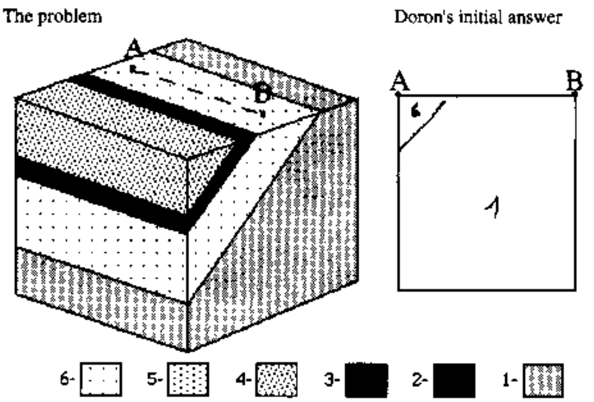
S: I can see only what is outside. How can I know what is inside?

*Figure 10.* An example of students' reasoning for nonpenetrative incorrect answers, in which the spatial configuration of the structure is not perceived by the student.

configuration of the structures. However, when they were not able to do so they gave incorrect answers which were based on analytic penetrative procedures. These findings suggest that the ability to perceive the spatial configuration of the structure is an obligatory factor, and must be involved. On the other hand, interviews of nonpenetrative students, who were not able to give correct answers because of their low VPA, even when they had a clear perception of the structure, indicate that VPA provides an obligatory factor as well.

Hence, the thinking processes of students in solving problems involving cross-sections of geologic structures are influenced by two obligatory and complimentary factors: (a) the ability to perceive the spatial configurations of the layers comprising a structure, and (b) the ability to penetrate visually into the image of the structure (VPA). These two factors can be classified under the category of spatial visualization as described by McGee (1979), and by Linn and Petersen (1985). These authors mention both the ability to create a mental image from a "pictorially presented object" and the ability to operate different mental manipulations of those images, as belonging to the category of spatial visualization.

Support for the model of the two complementary factors presented earlier can be found in case studies that were conducted with students working with the software Geo3D (Kali, 1993). This software is designed to assist high-school students in the perception of three-dimensional geologic structures, and provides manipulative animated visual illustrations of such structures as learning aids for solving problems. One type of illustration, layer disassembly, is designed to assist students in the perception of spatial configurations of geologic structures, and another type demonstrates "cutting" of the block diagram and revealing different cross-sections. The case studies, which were conducted as part of the evaluation of Geo3D, show different paths made by



S: *(picks up the form with the problem and looks at it from the left side, as if the structure was actually three-dimensional)* My drawing looks O.K..

I: Can you explain why?

S: Because if I'll take this part off *(points at the part of the block-diagram to the left of the cross-section)*, then these layers *(points at layers 1 and 6 at the right side of the right face of the block-diagram)* will continue straight inside...

I: Can you describe this structure?

S: It's like a wooden board with layers.

I: What do the layers look like?

S: They are flat, and they go this way *(puts his hands on top of each other, and tilts them to create an inclination)*.

I: Now imagine cutting those layers right through here *(interviewer shows the appropriate location of the cross-section on Doron's "hand-model")*.

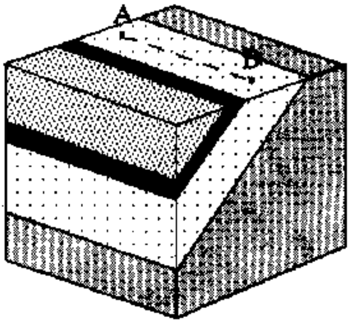
S: *(thinks a few seconds)* I think it will look very similar to my drawing but I'm not sure.

*Figure 11.* An example of students' reasoning for nonpenetrative incorrect answers, in which the spatial configuration of the structure is perceived by the student.

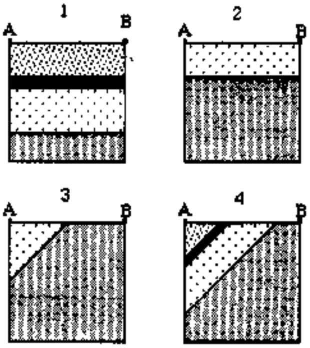
students throughout the software. An examination of these paths indicates that they remarkably fit the idea of the two complementary factors suggested. Students who displayed difficulties in the perception of the spatial configurations of the structures tended to manipulate the layer disassembly illustrations, and students who displayed difficulties in envisioning the cross-sections (low VPA students), tended to use the cutting animation.

Since the low VPA students in the model of two obligatory factors will always give incorrect answers, it is impossible to know the percentage of students who perceive the configuration of the structures, but who are not able to envision the cross-sections. However, the interviews provide evidence that low VPA students who gave incorrect answers to cross-section problems tended to identify the correct cross-sections when the problems were presented as multiple-choice questions, provided they were able to perceive the spatial configuration of the structure. The sample of the interviews is too small to offer generalizations, but presuming that this evidence is valid, it is suggested that a multiple-choice questionnaire would distinguish between low VPA students who are able to perceive the spatial configuration of the structure and low VPA students who are not able to perceive it. Moreover, the finding that multiple-choice problems assist low VPA students in solving cross-section problems has a practical implication:

The problem



multiple-choice possibilities:



S: It's this one *(immediately points on the second possibility, which is the correct answer)*, for sure! It's not this one *(points on the third possibility which is similar to his former answer)* that I drew before.

I: What made you change your mind?

S: It's clear to me now. I just got a little bit confused by these layers here *(points on the right side of the block-diagram)*, this is how they look from the outside, not from the inside.

*Figure 12.* An example of correct identification of a cross-section in a multiple-choice form of the problem, given by a student who initially gave a nonpenetrative incorrect answer.

It is likely that these students might improve their VPA by encountering appropriate multiple-choice problems. This finding supports Chadwick's (1978) assumption that students might improve their spatial skills needed for geology by encountering general spatial tests, which was later applied by Bezzi (1991) in his software.

Characterization of the abilities required in GeoSAT can also be accomplished by the analysis of score distributions. In many previous studies concerning spatial abilities, comparisons between different test items or between different groups were conducted via parametric statistical procedures. This indicates that the spatial abilities within the samples investigated were found or assumed to follow a normal distribution. The findings of the current study indicate that bimodal (U-shaped) distribution patterns are involved. These patterns were found even when results of males and females were analyzed separately. The existence of bimodal patterns suggests that perhaps the two factors mentioned earlier, which are required to solve the problems of GeoSAT, induce a certain mental barrier that is responsible for this almost dichotomous splitting of the population into two levels of success.

An examination of the frequency distributions reveals that in each subtest the bimodal pattern characterizes either the female or the male distribution, but it never describes the performance by both genders on the same subtest. Specifically, the bimodal pattern characterizes female performance on the cross-section and completion subtests and male performance on the construction subtest and on the entire test. An explanation of this phenomenon is based on two different findings: (a) the lower difficulty level of the cross-section and completion subtests compared with the construction subtest, as indicated in the validation and characterization of the instrument; and (b) the higher achievements of males compared with the females on the three subtests, which is in agreement with previous reports of gender-based differences in spatial abilities. Thus, it seems that bimodal patterns characterize both female and male score distributions, provided that appropriate difficulty levels of problems are involved. Accordingly, the easier subtests revealed bimodal patterns in the scores earned by females, while the males' scores tended toward the higher scores (negatively skewed); and conversely, the more difficult subtests revealed bimodal patterns in the male scores, while the females tended to perform poorly (positively skewed). These findings suggest that the barrier responsible for these bimodal patterns characterizes both females' and males' abilities to solve the problems of the test. However, the location of this barrier on an imaginary difficulty scale is at lower levels for the females and higher levels for the males.

A question that might be raised is whether this barrier exists in the ability to solve other spatial visualization problems, or whether it is a unique characteristic of spatial visualization skills needed in geologic perception. Previous studies characterizing spatial visualization of different populations did not mention bimodal distributions of performances (Battista et al., 1989; Ben-Chaim et al., 1988; Kyllonen, Lohman, & Snow, 1984; Russell-Gebbett, 1985; Small & Morton, 1983; Smith & Schroeder, 1981). Moreover, in many of these studies, parametric statistical procedures such as *t*-tests, and analyses of variance, which are based on normal distributions, were used.

Hence, it might be concluded that the skills needed to solve geologic problems of the type used in GeoSAT involve unique spatial visualization skills, which reveal the bimodal characteristic, and are not required in common spatial visualization tests. Such tests involve mental manipulations such as rotation (Ben-Chaim et al., 1986) or unfolding of objects (Bennet, Seashore, & Wesman, 1972), which refer to the external surfaces of the objects; the problems of GeoSAT refer to the internal parts, and thus involve the unique ability of VPA.

However, VPA is not a skill needed exclusively in earth sciences. Any subject matter involving problems that require envisioning shapes of intersections through three-dimensional objects should involve VPA. Examples from the technical domain are mechanical engineering, technical drawing, and architecture. In the scientific domain, many of the visualization problems require external manipulation and do not involve bisecting of objects. However, an example that involves internal spatial visualization is the perception of three-dimensional biologic structures, which are commonly presented through their intersections. Russell-Gebbett (1984, 1985) described two discrete skills used by secondary-school pupils, in solving problems involving three-dimensional structures in biology: "(a) the abstraction of sectional shapes, and (b) an appreciation of the spatial relationships of internal parts of a three-dimensional structure seen in differing sectional planes" (Russell-Gebbett, 1984, p. 223). The skills described by Russell-Gebbett seem to agree with the two factors involved in solving GeoSAT's problems. However, the barrier assumption cannot be examined in this example, since frequency distributions were not mentioned.

Considerable evidence exists for a dependent relationship between VPA and spatial visualization skills that refer to external surfaces of objects. Lord (1985) developed an instructional treatment program for enhancing the visuospatial aptitude of college undergraduates. The program included exercises that required the participants mentally to bisect three-dimensional geometric figures, and to envision the shape of the two-dimensional surface exposed by this bisection. This exercise resulted in significant improvement of the students' spatial orientation and spatial visualization.

Lord's rationale behind this program was that the main factor in the ability to envision the bisections is neural control of the image. This control determines the subject's ability to perform mental manipulations without losing the shape of the images. According to Lord, as long as the image can be held by the student the bisection is performed easily, while difficulties arise when the mental image is lost. However, the current study suggests that Lord's bisectional operations might correspond to the VPA, which itself is a factor that is independent of the neural control of an image. Accordingly, in the current study there were students who were able to hold images of structures without being able to bisect them. It is suggested, therefore, that the improvement in spatial orientation and spatial visualization of the students in Lord's study were induced not only by exercising imagery control, as indicated by Lord, but by exercising the VPA as well.

Finally, additional evidence for the relationship between the VPA and spatial visualization skills which refer to external properties of objects comes from the geology subject matter. Orion, Ben-Chaim, and Kali (1994) showed that 1st-year undergraduate students significantly improved their spatial visualization as a result of participation in introductory earth science courses. Spatial visualization was tested through tests that refer to external properties of objects. Some of the experiences mentioned by Orion et al. as inducing these improvements are geologic mapping at the field, the use of concrete models in the study of structural geology, and the study of fossil structure through their cross-sections in rocks. In light of the current study, these experiences might be considered VPA exercises. It is therefore suggested that the significant improvement in the performance of students reported by Orion et al. might indicate that the barrier which is perhaps involved in the VPA was crossed by many of the students during their geology education, and an improvement was induced in their external spatial visualization.

## Summary and Implications

1. Two complementary independent factors are required for solving GeoSAT's problems: (a) the ability to perceive the spatial configuration of the layers comprising a structure; and (b) the visual penetration ability (VPA), defined in the current study as the ability to envision internal cross-sections of structures.
2. Bimodal distributions characterize the scores achieved both by females and males in the three subtests of GeoSAT. These distributions suggest the existence of a barrier in the ability to solve the test's problems.
3. The performance of males was significantly higher than that of females, suggesting that males have a higher VPA than females.
4. Earth science students should be provided with appropriate assistance for enhancing their abilities to perceive and mentally bisect geologic structures. The assistance should be focused in two directions: (a) providing low VPA students with tools for envisioning cross-sections of structures, and (b) enhancing low VPA students' perception of spatial configurations of layers, especially those comprising complicated structures. Such assistance could be given by providing students with opportunities to disassemble models of geologic structures and investigate the spatial configuration of each layer comprising structures of different difficulty levels.

5. It is advisable to find those niches in science and technology education where VPA is needed and to provide all students, both males and females, with assistance in enhancing their abilities to perceive spatial structures of all kinds. The techniques for such assistance can be similar to those suggested for earth science students.

## Limitations and Suggestions for Further Research

One limitation of the study derives from the fact that VPA was notable only in students' answers to the four cross-section problems of the test. To obtain deeper validation of the existence of VPA, further research is suggested. Research based on questionnaires including more problems that require VPA might throw more light on this issue. In addition, such research might enable further examination of the suggestion that VPA is the cause for bimodal distributions. Another limitation derives from the small number of students interviewed. The outcomes of the interviews predict that multiple-choice questionnaires might enhance low VPA students, provided these students perceive the spatial configuration of the structures involved. Considering the significance of this outcome, it is suggested to design such a questionnaire and examine its effect on a larger number of students. Another point which was raised in the discussion and requires further examination is the relationship between internal and external visualizations. It will be interesting to determine whether exercising either of these skills may lead to an improvement in the other skill. Such knowledge is of great interest for the development of spatial training programs.

## References

Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex Roles*, *20*, 327–344.

Battista, M.T., Wheatley, G.H., & Talsma, G. (1989). Spatial visualization, formal reasoning, and geometric problem-solving strategies of preservice elementary teachers. *Focus on Learning Problems in Mathematics*, *11*, 17–30.

Ben-Chaim, D., Lappan, G., & Houang, R.T. (1985). Visualizing rectangular solids made of small cubes: Analyzing and effecting students' performance. *Educational Studies in Mathematics*, *16*, 389–409.

Ben-Chaim, D., Lappan, G., & Houang, R.T. (1986). Development and analysis of a spatial visualization test for middle school boys and girls. *Perceptual and Motor Skills*, *63*, 659–669.

Ben-Chaim, D., Lappan, G. & Houang, R.T. (1988). The effect of instruction on spatial visualization skills of middle school boys and girls. *American Educational Research Journal*, *25*, 51–71.

Bennet, G.K. Seashore, H.G., & Wesman, A.G. (1972). *Form T, differential aptitude tests, space relations*. New York: Psychological Corporation.

Bezzi, A. (1991). A Macintosh program for improving three-dimensional thinking. *Journal of Geological Education*, 39, 284–288.

Broadfoot, J.M. (1993, January). *Spatial ability and earth science learning*. Paper presented at the International conference on science education in developing countries: From Theory to Practice, Jerusalem.

Chadwick, P. (1978). Some aspects of the development of geological thinking. *Journal of Geology Teaching*, *3*, 142–148.

Cooper, M. (1992). Three-dimensional symmetry. *Educational Studies in Mathematics, 23*, 179–202.

Dyche, S., Mclurg, P., Stepans, J., & Veath, M.L. (1993). Questions and conjectures concerning models, misconceptions, and spatial ability. *School Science and Mathematics, 93*, 191–197.

Eliot, J. (1980). Classification of figural spatial tests. *Perceptual and Motor Skills, 51*, 847–851.

Hyde, J.S. (1990). Meta-analysis and the psychology of gender differences. *Journal of Women in Culture and Society, 16*,55–73.

Johnson, E.S., & Meade, A.C. (1987). Developmental patterns of spatial ability: an early sex difference. *Child Development, 58*, 725–739.

Kali, Y. (1993). *Research for characterization of the spatial abilities required in structural geology, and its implementation in the development of a software for high-school students.* Unpublished master's thesis, Weizmann institute of sciences, Rehovot, Israel.

Kiser, L. (1990). Interaction of spatial visualization with computer-enhanced and traditional presentations of linear absolute-value inequalities. *Journal of Computers in Mathematics and Science Teaching, 10*, 85–96.

Kyllonen, P.C., Lohman, D.F., & Snow R.E. (1984). Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology, 76*, 130–145.

Linn, M.C., & Petersen, A.C. (1985). Emergence an characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*, 1479–1498.

Lord, T.R. (1985). Enhancing the visuo-spatial aptitude of students. *Journal of Research in Science Teaching, 22*, 395–405.

Lord, T.R. (1987). A look at spatial abilities in undergraduate women science majors. *Journal of Research in Science Teaching, 24*, 757–767.

Maccoby, E.E, & Jacklin, C.N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

McGee, M.G. (1979). Human spatial abilities: psychometric studies and environmental, genetic, hormonal and neurological influences. *Psychological Bulletin, 86*, 889–918.

Orion, N., Ben-Chaim, D., & Kali, Y. (1994, March), *Relationship between earth science education and spatial visualization*. Paper presented at the National Association for Research in Science Teaching, Anaheim, CA.

Pearson, E.S., & Hartley, H.O. (1976). *Biometrika tables for statisticians* (3rd ed.). London, UK: Biometrika Trust.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space,* New York: Humana Press.

Rodriguez, W.E. (1990). A dual approach to engineering design visualization. *Engineering Design Graphics Journal, 54*, 36–43.

Russell-Gebbett J. (1984). Pupils' perceptions of three-dimensional structures in biology lessons. *Journal of Biological Education, 18*, 220–226.

Russell-Gebbett J. (1985). Skills and strategies—pupils' approaches to three-dimensional problems in biology. *Journal of Biological Education, 19*, 293–297.

Seddon, G.M., & Moore, R.G. (1986). The structure of abilities in visualizing the rotation of three-dimensional structures presented as models and diagrams. *Journal of Educational Psychology, 56*, 138–149.

Siegel, S., & Castellan, N.J. (1988). *Non-parametric statistics for the behavioral science.* (2nd ed.). New York: McGraw-Hill.

Small, M.Y., & Morton, M.E. (1983). Spatial visualization training improves performance in organic chemistry. *Journal of College Science Teaching*, *13*, 41–43.

Smith, W.S., & Schroeder, C.K. (1981). Preadolescents' learning and retention of a spatial visualization skill. *School Science and Mathematics*, *81*, 705–709.

Tuckey, H., & Selvaratnam, M. (1993).Studies involving three-dimensional visualization skills in chemistry: a review. *Studies in Science Education*, *21*, 99–121.

Wiley, S.E. (1990). Computer graphics and the development of visual perception in engineering graphics curricula. *Engineering Design Graphics Journal*, *54*, 39–43.

Received August 12, 1995
Revised October 31, 1995
Accepted November 16, 1995