JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 40, NO. 4, PP. 415–442 (2003)

Cognitive Factors Affecting Student Understanding of Geologic Time

Jeff Dodick, Nir Orion

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

Received 16 January 2002; Accepted 9 May 2002

Abstract: A critical element of the earth sciences is reconstructing geological structures and systems that have developed over time. A survey of the science education literature shows that there has been little attention given to this concept. In this study, we present a model, based on Montagnero's (1996) model of diachronic thinking, which describes how students reconstruct geological transformations over time. For geology, three schemes of diachronic thinking are relevant: 1. Transformation, which is a principle of change; in geology it is understood through actualistic thinking (the idea that present proceeses can be used to model the past). 2. Temporal organization, which defines the sequential order of a transformation; in geology it is based on the three-dimensional relationship among strata. 3. Interstage linkage, which is the connections between successive stages of a transformation; in geology it is based on both actualism and causal reasoning. Three specialized instruments were designed to determine the factors which influence reconstructive thinking: (a) the GeoTAT which tests diachronic thinking skills, (b) the TST which tests the relationship between spatial thinking and temporal thinking, and (c) the SFT which tests the influence of dimensional factors on temporal awareness. Based on the model constructed in this study we define the critical factors influencing reconstructive thinking: (a) the transformation scheme which influences the other diachronic schemes, (b) knowledge of geological processes, and (c) extracognitive factors. Among the students tested, there was a significant difference between Grade 9–12 students and Grade 7–8 students in their ability to reconstruct geological phenomena using diachronic thinking. This suggests that somewhere between Grades 7 and 8 it is possible to start teaching some of the logical principles used in geology to reconstruct geological structures. © 2003 Wiley Periodicals, Inc. J Res Sci Teach 40: 415–442, 2003

In anticipation of the curriculum reforms called for in Project 2061, a conference of educators and geoscientists was held in Washington, DC, in April 1988. At this conference participants identified those goals and concepts about planet Earth that every 17-year-old should know when completing his precollege education. Among the major concepts identified were ''The earth's natural processes take place over periods of time from billions of years to fractions of seconds'' (Mayer and Armstrong, 1990, p. 161).

This broad understanding of geological time influences many fields of science including evolutionary biology, cosmology, and ecology. In other words, if students are to understand and

Correspondence to: J. Dodick; E-mail: j-dodick@northwestern.edu
DOI 10.1002/tea.10083
Published online in Wiley InterScience (www.interscience.wiley.com).

even master such fields, they should first grasp their temporal framework. However, after reviewing the science education literature, Roseman (1992, p. 218) noted that there "was next to nothing about . . . how kids' understanding of notions of systems scale or models develops over time."

In a previous study, Dodick and Orion (2002b) reviewed the relevant science education literature on student understanding of geological time. They classified these studies into two broad groups: event-based studies and logic-based studies.

In event-based studies, students are probed about their understanding of the entirety of geological time (that is beginning with the formation of the earth or the universe). In other words, the researchers were primarily interested in exposing the students' alternative frameworks concerning temporally influenced phenomena and processes in earth history.

Major studies include Noonan-Pulling and Good's (1999) research on the understanding of the origins of earth and life among junior high school students and Trend's studies on the conception of geological time among 10- to 11-year-old children (Trend, 1997, 1998), primary teacher trainees (Trend, 2000), and inservice teacher trainees (Trend, 2001). In addition, it is possible to cite studies that included indirect or direct reference to geological time alongside other earth-related concepts (Happs, 1982; Schoon, 1989; Oversby, 1996; Marques & Thompson, 1997).

The research strategy of all these studies combined structured interviews with questionnaires in which subjects were asked to arrange major features of Earth history chronologically using either a physical timeline (Noonan-Pulling & Good, 1999) or a numerically divided time scale (Trend, 1997, 1998, 2000, 2001). In addition, Trend (1997, 1998, 2000, 2001) and Marques and Thompson (1997) interviewed subjects using pictures from different events from Earth history. The pictures acted as a stimulus for understanding student reasoning about chronology. Such picture-sorting tasks have a long tradition in psychological research on time conception [see, for example, Bullock, Gelman, & Baillargeon (1982), Friedman (1982), and Piaget (1969)]. In summary, all of these event-based studies reflect subjects' understanding of particular events and their serial order in geologic time.

This approach might be contrasted with the second type of research, the logic-based study. Logic-based studies focus on the cognitive processes undergone by students when solving problems involving geologic time. Within science education, Ault's (1981, 1982) pioneering investigation is the only substantial research of this kind. Based on the work of Zwart (1976), Ault (1981, 1982) defined time relationally. This definition implies that "time does not exist as a thing in itself, but rather is the coordination of relations among events and objects by an observer" (Ault, 1981, p. 7). Zwart (1976) suggested that the simplest relation for an observer to grasp is succession, or the before and after relationship, which in turn serves as the cognitive basis for understanding many different concepts of time, including order, simultaneity, and duration. Zwart's suggestion appears to have support from Stevenson and Pollitt (1986) who argued that children as young as 3 years of age understand the verbal meaning of "before and after." In contrast, Harner (1982) suggested that such understanding is acquired between ages 4 and 5.

Using this relational definition as a cognitive framework, Ault (1981, 1982) interviewed a series of students from kindergarten to Grade 6 (ages 5–11) and attempted to classify their answers on a series of problems testing time conception in general and geology specifically. Based on these interviews he concluded that a child's concept of conventional time was no impediment toward his understanding of geologic events. Indeed, many of the children in his test group were successful at solving puzzles involving skills necessary to an understanding of geological time, such as superposition and correlation. Nonetheless, in the environment of the field, these same children had difficulties in solving similar types of problems, indicating that there was little transfer from classroom work to authentic geological settings (Ault, 1982).

These difficulties can be traced to Ault's (1981) research design, which in turn was influenced by Piaget's (1969) work on time cognition. Piaget believed that a young child's understanding of time is tightly bound to the concept of motion, and many of his interview questions concerning time reflected this physics-based bias. Indeed, much of the work in cognitive and developmental psychology focuses on children's understanding of time and its relationship to motion.

Ault (1981, 1982) replicated some of Piaget's research methodology in his study. However, the science of geology builds its knowledge of time through the visual interpretation of static entities (formations, fossils, and landforms) (Frodeman, 1995, 1996). Ault's (1981) design involving motion simply multiplied the variables he needed to explain, as he admitted in a later work (Ault, 1982). Even if Ault's (1981, 1982) supposition that a young child's conventional time sense does not interfere with his understanding of the geological past is correct, this does not necessarily mean that he would able to understand geological time. This is because geological time possesses a special quality not seen in conventional time, an enormous scale which would likely complicate a child's thinking.

This argument is supported by work in developmental psychology. For example, Friedman (1982) suggested that between the ages of 6 and 8, the range of a child's temporal awareness is likely no more than a year. Moreover according to Friedman (1978) and Harner (1982), it is only at age 14 that children begin using time vocabulary and concepts such as century, generation, and forefather. Thus, it is unlikely that the children studied by Ault (1981, 1982) would have had a deep understanding of the immensity of geological time.

Not surprisingly, the widest body of research on time cognition comes from psychology. However, much of this work is concerned with conditioning or perception of time, topics with minimal applicability to education (Fraisse, 1982). However, within cognitive psychology, one area of inquiry, logical time (Fraisse, 1982; Friedman, 1978, 1990), has relevance to a consideration of geological time.

Friedman (1978) defined logical time as the cognitive ability making it possible to determine the temporal relations between events, which in turn permits one to construct a representation of these changes. In a broad sense this definition touches on aspects of geological time as well, and in fact this definition recapitulates the essence of Ault's (1981) research. By using geological strata as time units, one can relatively order events represented by such strata in time. Moreover, if one is lucky enough to find datable materials such as radioactive isotopes within such strata, they can provide an absolute time frame. Finally, one can reconstruct the environmental transformation of such strata through time.

It should be noted, however, that all pure psychological studies have tended to restrict themselves to scales of no more than years. This is an important difference because geological time, as we noted previously, adds a level of abstraction (its immense scale) which most people find difficult to conceptualize (Gould, 1987).

On a practical level, much of the cognitive research on logical time has concentrated on children no older than 11 years of age, the concrete operational stage. The research discussed in this article focused on students of junior high and high school age, during the formal operational stage; thus, the difficulties experienced by older children in understanding geological time may not be the same as those experienced by younger children in previous studies. Students in Israel as well as in many other countries encounter the earth sciences at the junior and senior high school levels; thus, this study has a practical element in that it will help guide future teaching practice.

Deep Time: Suggested Cognitive Model

Based on previous research in both science education and cognitive science, it became apparent to researchers that an understanding of geological time might be broken down into two different concepts. [This definition was first formulated by Dodick and Orion (2002b) and is replicated here for purposes of clarity.]

1. A passive temporal framework in which large-scale geological and biotic events occur. It is suggested that this definition of deep time depends on the connections people build between events and time. In the cognitive literature this is comparable to Friedman's (1982, 1990) associative networks, a system of temporal processing well-suited for storing information on points in time: for example, associating the holidays with seasons or weather. By this reasoning, this understanding of deep time should be mitigated by a person's knowledge of such events. In essence, this definition is closer to the first type of science education research mentioned above, which was based on event-based knowledge.
2. An active logical understanding of geological time used to reconstruct past environments and organisms based on a series of scientific principles. In essence, this definition is similar to the second type of science education study noted above, which is based on the logical understanding of geological time. It also reflects Friedman's (1978) definition of logical time mentioned previously, with the proviso that Friedman was concerned with human understanding of conventional time systems.

Based on the second concept, it might seem that students untutored in the basic logic of geology would be unable to reconstruct the changes which have affected a fossil sequence or a depositional system; however, the structure of geological logic is comparable to Montagnero's (1992, 1996) model of diachronic thinking. Montagnero (1992, 1996) defined diachronic thinking as the capacity to represent transformations over time; such thinking is activated, for example, when a child attempts to reconstruct the growth and decay of a tree. Of course, trees are a natural part of any child's environment, so it was considered to be a particularily relevant item for Montagnero (1996) to investigate. However, it is argued here that his concept of diachronic thinking might be equally applicable to more specialized scenarios such as the deposition through time of many geological structures.

Montagnero (1996) built a model which defines the structural and functional entities that are activated when diachronic thinking is employed. These consist of four schemes which permit a subject to understand transformations over time. As part of this research, these schemes have been translated into the specific logical skills needed to solve temporal problems involving geological strata (Table 1).

According to Montagnero (1992, 1996), children as young as 10–11 years of age are able to activate a full diachronic perspective based on the four schemes outlined in Table 1. The limiting factor to this ability is the subject's knowledge of a given phenomenon. Based on structured interviews with children between the ages of 7 and 11, Montagnero (1996) denoted three sources of knowledge at the origin of the diachronic schemes.

1. Empirical knowledge: knowledge of transformations derived from personal experience as well as through the influence of specific cultural representations.
2. Organizational knowledge: includes an understanding of dimensions (numbers, space, and time) as well as causal relations.
3. Axiological knowledge: knowledge of transformations based on the subjects' own system of values. Thus, for young children, systems are often judged as progressing

Table 1
Diachronic schemes and their geological correlates

Diachronic Schemes and Their Explanation (Based on Montagnero, 1996)	Geological Correlate
Transformation: This scheme defines a principle of change, whether qualitative or quantitative. Quantitative transformation implies an increase or a decrease in the number of elements comprising an object—for example, the changing number of leaves on a tree during different seasons. Qualitative transformations are concerned amongst other things with the complexity of objects, such as the change in shape of a growing tree	In geology, such changes are understood through the principle of ''actualism'' (''the present as key to past'') in which geological or biological change is reconstructed through comparison with contemporary fossil and depositional environments
Temporal Organization: This scheme defines the sequential order of stages in an evolutive (or transformational) process. It also provides the general form of the sequence of stages for example linear, cyclical, etc.	In geology, logical principles, including superposition, correlation, and original horizontality, all of which are based on the three dimensional relationship among strata are used as a means of determining temporal organization
Interstage Linkage: The connections between the successive stages of evolutive phenomena. Such connections are built in one of two ways: 1. Relations between necessary prerequisite and its sequel. 2. Cause and effect relationships	In geology such stages are reconstructed via the combination of actualism (as defined above), and (scientific) causal reasoning
Dynamic Synthesis: Forming a whole from a set of successive stages which are thus conceived of as a manifestations of a single process of change	In geology such a dynamic synthesis is not a separate scheme but is rather a result of correctly activating the principles discussed previously. For this reason, this element was not emphasized in this study

toward or away from a preferred state (for example, a tree that is slowly decaying as it grows older).

The first two sources of knowledge are of prime importance to any understanding of geological transformation. In terms of empirical knowledge, if the subject does not know, for example, that limestone forms in shallow water, he may not be able to reconstruct the stages in which a series of limestone strata formed. In this study, we attempted to limit the types of knowledge which might interfere with a subject's ability to represent geological transformations over time. This was done by providing the information directly, or by using clues that all of the subjects understood (for example, clams live in the sea, implying a former marine depositional environment).

In terms of organizational knowledge, a prime example from geology is the relationship between the proportional size and number of layers and their supposed age. Experienced geoscientists understand that size and number of layers do not necessarily translate into proportional amounts of time passed. As will be seen, this understanding was not held by most of the students sampled.

In addition to the knowledge factors which might interfere with diachronic thinking, one should also give attention to the innate cognitive abilities of the subjects tested. For example, geological strata are three-dimensional structures that require visual interpretation before one can draw conclusions concerning their physical attributes (Chadwick, 1978; Kali & Orion, 1996).

Purpose

The purpose of this study was to define the factors which affect a student's ability to reconstruct geological systems, such as depositional environments and fossil sequences, that have developed over time. In this way, we hoped to expose the alternative frameworks possessed by the subjects in understanding the concept of geologic time as expressed in stratigraphic structure.

Using Montagnero's (1996) model of diachronic thinking as a guiding framework, we focused on the logical strategies students employ when confronted by temporal problems in the earth sciences. Thus, we examined both the way students employed the diachronic schemes as well as the possible factors (knowledge and innate cognitive abilities such as spatial visualization) which might interfere with this process. In this way it was possible to construct a cognitive model which might account for how students reconstruct geological structures.

This study was conducted with a population of students from Grades 7–12; thus, a concurrent goal of this research was to determine how such diachronic schemes develop in students of different age. On the practical level of education research, this aspect of the study is significant because it permits us to make recommendations about when it might be possible to begin teaching various elements of historical geology.

Method

To expose the alternative frameworks and logical strategies employed by students when confronted by temporal problems in geology, we used a combination of quantitative and qualitative research methods; these included written questionnaires, structured interviews, and observations of students in class.

The questionnaires employed in this research were all designed specifically for this study and included the Geological Time Aptitude Test (GeoTAT), the Temporal Spatial Test (TST), and the Strategic Factors Test (SFT). They will be described in detail in the next section of this report. The questionnaires shared a common feature in that they all consisted of a series of open pictorial puzzles permitting one to classify both the subjects' answers, as well as the strategies they employed in answering these puzzles. This open format was chosen because it combined some of the features of the structured interview with the ability to build a statistical sample. In this way it was possible to make generalizations about the alternative frameworks held by subjects in this sample as well their ability to solve temporally based problems in the earth sciences. Furthermore, it permitted us to build a detailed age-based profile of the research sample.

Instruments

GeoTAT. The GeoTAT consists of six multipart open geological puzzles which tested the subjects' logical understanding of the temporal relationship among geological strata and their fossil contents (see Appendix).

A limitation of written instruments is that students who have reading difficulties sometimes receive lower scores because they do not understand the meaning of the questions, rather than as a result of lower cognitive abilities. To limit this constraint, the GeoTAT was administered to the subjects by the first author during a standard 45-minute classroom period. He remained present during its writing and encouraged the students to query him any time they were not sure about the puzzles' wording, so that it could be explained. Another advantage to his involvement in this process was that it permitted him to gauge the level of effort invested by the subjects.

Additional insight into the subjects' understanding was provided by comments they made during the writing of the GeoTAT, as well as by a limited number of in-depth interviews held with them upon completion of the test.

A third source of information was gleaned from a second sample of students who were studying a unit in fossil biology, "From Dinosaurs to Darwin" (Dodick & Orion, 2002a), which emphasized the temporally mediated changes that have affected the earth's environment and its biota. These students were tested with the GeoTAT both before and after completing the program; they were also monitored and interviewed by the first author while they worked on this program's various investigations.

Table 2 outlines the geological and corresponding cognitive skills tested in each puzzle of the GeoTAT. In this table, *Geological skills* refers to the principles and techniques a geologist or advanced geological student would use to reconstruct geological structures and systems temporally. *Cognitive skills* refers to the diachronic schemes correlated with these geological principles.

Based on a factor analytic study presented in Dodick and Orion (2002b), the GeoTAT puzzles were distributed into the following three groups.

1. Isolated transformation puzzles. These puzzles (6a and 6b) require the subject to use only the diachronic scheme of transformation. This corresponds to a geologist's use of actualistic thinking.
2. Isolated temporal organization puzzles. These puzzles (1a, 4, and 5) require activation of only the temporal organization scheme. In geological terms, this means the use of principles such as superposition, correlation, and bracketing.

Table 2
Geological skills and their corresponding cognitive skills tested in the GeoTAT

Puzzle No.	Puzzle Identification	Geological Skill(s)	Cognitive Skill(s) (Diachronic Schemes)
6a	Transformation I	Actualistic thinking	Isolated transformation
6b	Transformation II	Actualistic thinking	
1a	Temporal Organization I	Superposition, lateral horizontality	Isolated temporal organization
4	Temporal Organization II	Superposition, bracketing	
5	Temporal Organization III	Superposition, correlation	
1b	Diachronic I	Superposition, lateral horizontality, actualistic thinking, and causal thinking	Full diachronic schemes (transformation, temporal organization, and intestage linkage)
2	Diachronic II	Superposition, actualistic thinking, and causal thinking	
3a	Diachronic III	Superposition, actualistic thinking, and causal thinking	
3b	Diachronic IV	Superposition, actualistic thinking, and causal thinking	
6c	Diachronic V	Superposition, actualistic thinking, and causal thinking	

3. Full diachronic schemes. These puzzles (1b, 2, 3a, 3b, and 6c) require the subject to use the full complement of diachronic schemes noted in Table 1 (temporal organization, transformation, and interstage linkage). Geologically, this means use of a combination of skills corresponding to each of the diachronic schemes (superposition, lateral horizontality, actualistic thinking, and causal thinking).

In each of the three delineated groups there are a variety of different puzzles which emphasize different shadings of the same diachronic schemes. Such task variation, a principle feature of cognitive research (Friedman, 1982), provides important information about the nature of competence underlying the subjects' ability to activate the diachronic schemes. Many studies of temporal cognition published up to the 1960s were mostly descriptive in nature and lacked any such task analysis, which made their results unreliable.

The GeoTAT was evaluated in two ways.

1. The puzzles requiring the full complement of diachronic skills were scored by determining whether they conformed to the three schemes of Montagnero (1996). Thus, points were allocated based on the presence of temporal organization, transformation, and interstage linkage. No points were granted for dynamic synthesis because in geology such a dynamic synthesis is not a separate scheme but rather is a result of correctly activating the other schemes (Table 1).
2. The remaining puzzles were marked according to how closely they fit the correct geobiological scenarios designed by the researcher. Two other researchers in science education validated this fit between puzzle and answer (Dodick & Orion, 2002b).

As with all studies, the GeoTAT profile generated a host of new research questions. To better answer these questions, two other instruments were developed for this study: the TST and the SFT, which are described below.

TST. This test is divided into two parts.

1. Four selected puzzles from the GeoTAT questionnaire (Temporal Organization I, Temporal Organization III, Diachronic II, and Diachronic V).
2. Fourteen puzzles from the Middle Grades Mathematics Project (MGMP) Spatial Visualization Test. Originally consisting of 32 multiple choice puzzles, this statistically validated questionnaire tested the students ability to manipulate two-dimensional figures (building maps) and three-dimensional solids (cubes) visually (Ben-Chaim, Lappan, & Houang, 1986, 1988). The three-dimensional puzzles used in this study were chosen with the aid of David Ben-Chaim, one of the designers of the MGMP spatial visualization test.

This instrument was distributed during a standard 45-minute class to a sample of Grade 10 and 11 students with no background in geology. As with the GeoTAT, it was monitored by the first author.

SFT. This test consists of three illustrations representing pairs of three-dimensional outcrops with differences in overall size and number of layers. The subjects were required to estimate, if possible which of the paired outcrops was older. The subjects were provided with four possible answers: Outcrop A, Outcrop B, no differences in age, and impossible to determine age. After each set of puzzles, subjects were asked to explain their reasoning for the answer they chose.

The SFT required 25 minutes to complete and was distributed to a sample of Grade 11 and 12 students majoring in geology. The structure of this test was based on a series of five in-depth interviews held with geology students in Grades 9, 11, and 12, which tested the proposed format of the puzzles.

Sample

Each of three questionnaires used in this study was distributed to different samples. Each will be discussed in turn.

The GeoTAT sample consisted of 285 Grade 7–12 students selected from 15 classes in five cojoined junior and senior high schools from urban areas in Israel. This sample is further broken down in Table 3.

The age-based distribution of subjects between Grades 7 and 12 was mandated by two factors: (a) It permitted the authors to extend the range of Ault's (1981, 1982) study, which concentrated on children from Grades K–6; and (b) students in both junior (Grade 7–9) and senior (Grade 10–12) high schools in Israel now have the opportunity to study earth science topics in school. This research allowed the authors to make recommendations concerning the future content of the earth science program.

No subjects in this research sample were studying geology as part of their matriculation studies.

The TST was distributed to a sample of 172 Grade 10–11 students from seven classes in five different high schools (Table 3). None of the subjects had a background in the earth sciences.

The SFT was distributed to 52 Grade 11–12 high school earth science majors from three different classes in three different schools who study geology as part of their matriculation exams (Table 3). This sample was chosen to control factors of content knowledge, specifically geological deposition. This decreased the chances that student difficulty with size and number factors was not connected to their lack of knowledge of basic geological processes. Moreover, if the earth science students had difficulty with such factors, it is probable that nongeology majors would find such problems even more difficult. In other words, the earth science students provided an upper baseline of achievement.

Data Analysis

For the GeoTAT, results from individual classes were pooled together by grade (Grades 7–12). As many of the puzzles carried a different raw score, they were standardized to a mark of 100% for easier comparison. They will be reported as such here. Mean scores, standard deviation, and sample size were determined for each of the grades (7–12) tested (Table 4). One-way analysis of variance (ANOVA) was used to determine whether there was a difference among the grade mean scores for each of the GeoTAT puzzles. Moreover, to determine how each of the six grades

Table 3
Breakdown of questionnaire samples

Questionnaire	Grades	*N* (Classes)	*N* (Sample)	*N* (Male)	*N* (Female)
GeoTAT	7–12	15	285	130	155
TST	10–11	7	172	74	98
SFT	11–12	3	52	21	31

Table 4
Analysis of variance comparing grade mean scores of results obtained on GeoTAT puzzles

Puzzle	12		11		10		9		8		7		F	p	Duncan ($p < .05$)
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*			
1a	53.7	33.9	78.0	22.7	59.8	29.8	57.6	35.7	41.1	28.6	47.1	32.2	7.4	.001	11 > 10, 9, 12, 7, 8; 10, 9 > 8
1b	17.6	25.5	18.2	25.5	9.8	16.7	8.9	12.9	5.9	14.0	4.4	9.6	4.6	.001	11, 12 > 10, 9, 8, 7
2	35.9	30.7	25.9	25.0	34.6	31.3	24.3	28.5	16.2	22.3	15.5	19.5	4.9	.001	12, 10 > 8, 7
3a	43.5	30.4	32.4	31.2	34.2	30.0	36.4	29.5	12.8	19.4	15.0	16.8	9.8	.001	12, 9, 10, 11 > 7, 8
3b	49.1	39.4	44.3	39.2	43.9	40.6	38.1	38.7	17.0	31.8	16.3	26.3	7.1	.001	12, 11, 10, 9 > 8, 7
5	55.0	32.8	44.5	38.1	44.6	36.0	46.6	32.9	27.9	33.9	20.3	30.6	6.7	.001	12, 9, 10, 11 > 8, 7
6a	83.3	32.2	76.1	41.1	67.1	44.2	67.8	43.3	42.6	47.8	21.3	33.7	14.5	.001	12, 11, 9, 10 > 8, 7; 8 > 7
6b	48.2	50.4	52.3	50.5	36.6	48.8	39.0	49.2	17.0	38.0	12.5	33.5	5.6	.001	11, 12, 9, 10 > 8, 7
6c	54.1	35.5	43.6	37.7	39.5	36.5	38.0	39.8	16.2	28.8	13.0	25.0	10.0	.001	12 > 9, 8, 7; 11, 10, 9 > 8, 7

differed specifically for each of the puzzles, Duncan's new multiple range test was also used (Huck, Cormier, & Bounds, 1974). All differences in the Duncan's test were evaluated at a significance level of .05 (Table 4). In presenting Duncan's new multiple range test (Table 4), the mean grade scores were arranged in order on the basis of the size of the means. Moreover, all grade levels separated by commas in the table are not significantly different. All grade levels that come before a greater than sign in the table are significantly larger than all of those grade levels which come after this sign.

In the TST questionnnaire, mean scores were determined for both the four selected GeoTAT and 14 MGMP puzzles using the entire sample as well as the individual classes. Product moment correlation coefficients were determined for both the entire questionnaire (GeoTAT vs. MGMP) as well the individual GeoTAT puzzles against the 14 puzzles of the MGMP.

In the SFT questionnaire, cumulative scores were determined for the four possible answers of each puzzle. In addition, both the explanations of the incorrect and correct answers were classified to provide better insight into the students' understanding of geological structures and their relationship in time.

Results

Table 4 presents the ANOVA for each puzzle of the GeoTAT as well as Duncan's new multiple range test of the results on the GeoTAT puzzles.

Based on analysis of the GeoTAT results, a number of general trends could be discerned.

1. In 6 of 10 puzzles there was a significant difference ($p < .05$) between Grade 9–12 students and Grade 7–8 students. This trend is found in Puzzles 6a (Transformation I), 6b (Transformation II), 5 (Temporal Organization III), 3a (Diachronic III), 3b (Diachronic IV), and 6c (Diachronic V).
2. In 6 of 10 puzzles there was no significant difference among Grade 9–12 students; this includes 5 of the puzzles noted above (Puzzles 6a, 6b, 5, 3a, and 3b) as well as Puzzle 4 (Temporal Organization II).
3. When we break down the results of the five full diachronic puzzles, we note that among the three diachronic schemes, the temporal organization scheme consistently received the highest mean scores at each grade level. In other words, most students found this scheme easiest to activate in these puzzles. However, with the exception of Puzzle 2 (Diachronic I) the strata in these puzzles were neither tilted nor folded.

The following is a more detailed analysis of the results according to the tripartite classification of this instrument first presented in the factor analytic study of Dodick and Orion (2002b) and replicated in Table 3. In each of the three sections the results are analyzed according to puzzle content and the specific elements of diachronic thinking which posed difficulty for the sample tested in this study.

Puzzles That Tested the Isolated Transformation Scheme

Transformation I (Puzzle 6a). In this puzzle students were provided with a schematic map of a dinosaur excavation site. Two areas were delineated on this map: Area A consisted of an exposure of terrestrially deposited sedimentary rock containing the fossils of dinosaurs (from babies to adults) and their nests; Area B consisted of marine deposited sedimentary rock containing fish fossils. The subjects' task was to reconstruct this site. To do so, students had to draw on an understanding of the historical connection between present and past conditions, a technique known as actualism in geology, and associated in this study with the transformation scheme of diachronic thinking. If the proper parallels are drawn between the past and present, subjects should arrive at the conclusion that Area A was an island surrounded by Area B, a low-lying sea. Most students found this puzzle easier to do than any other puzzle (Table 4).

Transformation II (Puzzle 6b). The second isolated transformation puzzle required the subject to interpret the significance of the dinosaur population structure found on the island. In information provided in the previous puzzle (Transformation I Puzzle 6a), it was noted that the dinosaurs found at this site varied greatly in size from newly hatched babies to full-grown adults. This suggested that these dinosaurs lived in family groups and that they were killed in a sudden, fast instinction episode.

In general, this puzzle was far more difficult than the previous one (Transformation I Puzzle 6a) as it relied on the subject's ability to piece together family structure or sudden extinction events; the corresponding mean scores were therefore lower. Similar to the previous puzzle, there were significant differences between students in Grade 9–12 and those in Grades 7–8. Moreover, there was also a significant difference between Grade 8 and 7 subjects.

As with the previous puzzle, an understanding of the pictorial and visual details was key to this puzzle's solution; moreover, a pilot study using this puzzle indicates that biological knowledge is a factor influencing this puzzle (Dodick & Orion, 2002b).

The most prominent alternative conception was that the evidence indicated that "the dinosaurs were living in Area A for a long time." However, this answer was predicated on the understanding that this was a permanent nesting site permitting the rearing of many generations; however, no evidence is provided in this puzzle which might support this supposition. Moreover, it discounts the possibility that the growth rates of the dinosaurs were fast. (In fact there is much scientific evidence to support this conclusion.) In other words, a correct solution is also dependent on filtering the data provided so as not to assume solutions not supported by the evidence. This ability to know the limits of one's evidence is key when dealing with indirect historical data.

Puzzles That Tested the Isolated Scheme of Temporal Organization

Puzzle 1a (Temporal Organization I). In this puzzle, the subjects were presented with a picture of a folded rock exposure containing fossiliferous layers; their task was to order the fossils temporally from oldest to youngest. To complete this task successfully, subjects needed to order the fossils based on the relative position of the layers in which they were found.

The relative position of the layers is emphasized to differentiate it from a situation of simple superposition in which stratigraphic layers are undisturbed (i.e., not tilted or folded). In such a case, the age of the fossils corresponds to the relative position of the layers as well as the relative height of the layers. However, to employ such an understanding with disturbed layers often leads to an incorrect age interpretation. Thus, in this puzzle the folded nature of the layers is such that there are fossils which sit relatively higher than their neighbors but are in fact are positioned in layers which were deposited below their neighbors, and therefore should be interpreted as being older.

That the relative position of the layers determines relative age is better understood if one activates a second geological rule, original horizontality, the principle that all sedimentary bedding is horizontal at the time of deposition (Press & Siever, 1982). Thus, if the bedding is currently not horizontal, it can be assumed that it was disturbed by some natural force. It was not expected that the students would formally recognize this principle, but unconsciously they might be able to use it by temporally ordering the fossils according to the strata in which they were found rather than by using the misleading clue of relative height.

Based on the structure of this puzzle, students who employed a strategy of original horizontality in this puzzle received a minimum score of 50%. The results for this puzzle indicate that Grade 7–8 classes were less likely to use this strategy, scoring a maximum mean score of 47.1%.

Surpisingly, Grade 12 students had almost as many difficulties as the Grade 7–8 sample; statistically there was no significant difference between these two groups (Table 4). Indeed, in the next puzzle (Temporal Organization V, Puzzle 5), which relies in part on the use of the principle of superposition, the Grade 12 students received higher mean scores than all of the other subjects (Table 4). This suggests that much of the difficulty experienced by the Grade 12 group in this puzzle was connected to their difficulty in activating the principle of original horizontality.

These results also suggest that the Grade 9–11 students were developing an understanding of the temporally mediated changes influencing these strata, because they were able to visualize the changes from horizontal to folded layers that had to be inferred from the present condition to solve this problem. This is corroborated by the fact that they performed significantly better than the Grade 7–8 students in all of the full diachronic puzzles, as will be shown.

Question 4 (Transformation II). The second puzzle in this group is a three-dimensional rendering of an outcrop consisting of three thick layers of fossiliferous sedimentary rock. The middle sedimentary layer is bounded above and below by layers of igneous rock which have been dated radiometrically using isotopes to 100 million years (lower) and 90 million years (upper) in age, respectively. Based on this information, the student is required to state the age of the fossils in each of the sedimentary layers. In fact, the correct technique is to say that age of the sedimentary layer is between the measured ages of the intervening igneous layers. Thus, in this puzzle the correct age of the middle layer and the fossils it contains was "between 90 million and 100 million years" in age. The lower layer was >100 million years in age, and the upper layer <90 million years in age. In other words, this answer was obtained by bracketing the sedimentary exposures above and below, by igneous rock layers, in which absolute age is measured.

For the vast majority (98%) of students this puzzle proved to be too difficult to solve. Indeed, this puzzle posed problems even for geology majors: The mean score for the 36 Grade 12 students tested with this puzzle was 33.0%, whereas the mean score for the 22 Grade 11 students was only 14.2%.

Interestingly, the same mistake continually appeared among all subjects who did not solve this puzzle correctly, apportioning equal amounts of time (in this case 10 million years) to each of the three sedimentary layers in the puzzle. Each of the sedimentary layers in this puzzle was of the same thickness (see Appendix Puzzle 4). This suggested the possibility that the test subjects believed that the absolute age of the layers was connected to its proportional size. (This assumption was tested by the SFT.)

Puzzle 5 (Temporal Organization III). In this puzzle, students were presented with pictures of three separated outcrops, each consisting of five layers containing fossils. The task was to order temporally, from oldest to youngest, the entire set of fossils represented in these three outcrops. To do so, students needed to use a technique known as stratigraphic correlation. This technique is based on the principle that two geological layers, which are geographically separated, are of the same age if one finds preserved within them the same type of fossil organism (Montgomery, 1993). Thus, the key to this puzzle was the ability to match identical fossils which were found in more than one layer, while recognizing the fact that the layers in which they were found were of the same age.

Because it introduces a second level of abstraction, synchrony in time, correlation is more complex than superposition which principally depends on the before and after relationship. Thus, although the principle of superposition had a role in this puzzle, its actual use could lead to mistakes if the subject equated each fossil found at the same level of strata, but separated geographically, as being of the same age. In fact, the proper technique is to ignore such false clues and instead to integrate the synchronous strata from the separated stratigraphic columns. Such integration requires a three-dimensional strategy that incorporates both vertical movement along the outcrops as well as translation between outcrops.

Based on the structure of the puzzle, subjects who restricted themselves to a strategy employing superposition could achieve a maximum score of only 40%. In contrast, a score of 50% or better indicates that students were beginning to use the strategy of stratigraphic correlation. Among the research sample, mean scores obtained by Grade 7 and 8 students indicated that at best they usually favored a strategy of superposition (Table 4). In fact, many of these students randomly ordered the fossils, as they even had difficulty activating the principle of superposition in this complex puzzle. Indeed, it was not rare to see such students order three different fossils from three different localities temporally equivalent because they appeared on the same stratigraphic level. This is the primary reason that in terms of overall trends there were significant differences between subjects in Grades 9–12 and subjects in Grades 7–8.

Puzzles Requiring Full Diachronic Puzzles

Puzzle 1b (Diachronic I). This puzzle is a continuation of Transformation I (Puzzle 1a). Here, however, subjects were required to reconstruct the events which created the folded structure of the fossiliferous rock exposure in Transformation I. A full answer should include the following elements in this order: deposition of strata, uplift, folding, and erosion.

The mean scores were low for all age groups, with significant differences between subjects in Grades 11 and 12 and those in Grades 7, 8, 9, and 10, respectively (Table 4). The low mean scores reflect the fact that most students related only to the first event of this sequence: deposition of the strata (Table 4). Few arrived at the next stage of reconstruction: folding. In part this is due to a lack of knowledge concerning:

1. The significance of folded strata. Even though they were provided with the clue that marine sedimentary rock is originally deposited in horizontal layers, they treated the folding as the original orientation of the rocks. In other words, they were unable to employ the transformational scheme of diachronic thinking so as to visualize the structure of the layers before folding.
2. The causal agent for such folding, such as localized (tectonic) pressure. In other words, the interstage-linking scheme of diachronic thinking.

In summary, it might be said that this problem was difficult for all of the students because it was composed of a series of stages (deposition, uplift, folding, and erosion) each requiring the activation of the full suite of diachronic schemes. However, as the clues provided did not elicit these schemes, the students had major difficulty in solving this problem.

Puzzle 2 (Diachronic II). In this puzzle, the subjects were required to reconstruct the events which formed an outcrop consisting of alternating layers of fossil-bearing marine and terrestrial sedimentary rock. Answered correctly, the subjects should have concluded that they were ultimately formed by a cyclical change in sealevel (termed a transgressive–regressive sequence in geology). The key to solving this puzzle was in connecting the fossils and rock types to the environments in which they were deposited. To do this, the subjects needed to employ actualistic thinking, sometimes defined in geology as the "the present as key to the past," or in terms of diachronic thinking the transformation scheme (Table 1). Indeed, when the mean scores were broken down into their constituent diachronic schemes, there was an extremely close fit between the mean scores received on the transformation scheme and the results for the entire puzzle.

The critical nature of the transformation scheme is also attested to by the fact that its correct application also affected the next diachronic scheme, interstage linkage, which is associated with cause and effect thinking. This is because, once the subject correctly understood the change in environment through time, he had a better chance of understanding that it was driven by temporal changes in sea level. Indeed, if we compare the mean scores for the two diachronic schemes we find that they are similar.

Overall, there were significant differences between students in Grades 10 and 12 and those in Grades 7 and 8, respectively (Table 4). Students in Grades 11 and 9 formed a zone of transition between these groups.

Puzzle 3a (Diachronic III). In this puzzle subjects are presented with an illustration of an outcrop containing a sequence of fossil horse toes and are asked to provide a logical scenario. To achieve a full score, subjects need to reconstruct this outcrop as containing an evolutionary sequence in which the fossils gradually transform through time from the primitive (four-toed) condition to the advanced (one-toed condition). Complicating the interpretation of this puzzle was an assumed conflict between two opposing arrows of temporal organization—geological and evolutionary—represented by this outcrop. The geological arrow is correctly resolved by applying the principle of superposition. The direction of the evolutionary arrow may also be inferred by this same principle. In other words, if a fossil bearing exposure is undisturbed, any evolutionary sequence contained within should parallel the depositional sequence of strata (from lower to higher). Thus, a correct interpretation of the evolutionary sequence in this puzzle is indeed a loss of toes as an adaptation to changes in the environment.

However, a number of students in all grades suggested that the correct evolutionary sequence was in fact from a single digit to many digits, which directly conflicts with the geological arrow of time. This is based on the classic alternative framework that evolution progresses toward greater complexity, and for some of the subjects participating in this research such complexity is represented by a larger number of bones; in this case, toes.

In terms of diachronic thinking, this alternative framework of evolution interferes with students' activation of the transformation scheme, which in turn influences students' understanding of the temporal organization scheme creating conflict with the true direction of temporal organization. This alternative framework has been encountered by the authors in high school, university, and INSET courses. Indeed, during a preliminary phase in this research, in a series of structured interviews, both junior and senior high school students consistently echoed this view of evolution.

In this study, this alternative framework was most expressed among Grade 7 and 8 students who achieved a mean score of no greater than 45.0% for the scheme of temporal organization. When one considers that the actual temporal organization is provided by simple superposition, in that the strata are neither tilted nor folded, this mean score can be said to be low. Possibly this alternative framework may be influenced by a faulty analogy to technology, wherein new technological devices are often considered to be more complex than their predecessors, in part because they contain a larger number of parts. On a practical level, for those interested in teaching evolution it is important to make students aware that the loss and fusion of bones is often an advanced condition among vertebrates.

Another area in which the transformation scheme indirectly affected the results was in providing the causal reason (i.e., the scheme of interstage linkage) for the change in the skeletal structure, which is evolution. Some of the subjects suggested instead that this sequence represented a slow deterioration of the fossil content. This answer suggests that the subjects did temporally organize the fossil sequence correctly, but for the wrong reasons. They did so because they interpreted the fossil sequence as they might interpret a time-lapsed sequence such as in a film, in which deterioration slowly removes organic material. In other words, they believed that the deterioration was a dynamic process occurring within the static layers of rock.

With so many difficulties connected to the scheme of transformation, it is not surprising that the highest mean score achieved on this puzzle was no greater than the 43.5% of the Grade 12 subjects.

Question 3b (Diachronic IV). This puzzle focused on the strata of the upper half of the outcrop presented in the previous puzzle (Diachronic III, Puzzle 3a). In contrast to the strata in the lower half, which contain an evolutionary sequence, the upper half is devoid of fossils. This puzzle required the subject to determine why the fossil sequence in the lower half of the outcrop disappears in the upper reaches of the stratigraphic column. To answer this puzzle correctly, students had to consider external causal factors which might eliminate the fossils. For example, a change in the environment might cause local extinction and/or geographical displacement of the population. In other words, in this puzzle the diachronic scheme of interstage linkage was the focus.

The major source of difficulties was misinterpretation of the transformation scheme, which was carried over from the previous puzzle (Puzzle 3a). This is supported by the fact that the mean scores in this puzzle were close to the scores achieved in the previous puzzle (3a). Moreover like the previous puzzle, there were major significant differences between subjects in Grades 9–12 and those in Grades 7–8, respectively.

More important, many of the individual answers reflected the same alternative frameworks regarding the scheme of transformation. Thus, the most common alternative frameworks noted were that the fossil bones continued their deterioration until nothing existed, or that the organisms represented by the fossils literally lost the entire limb and thus became extinct through the evolutionary process. The problems with the former conclusion have already been discussed. The latter solution is in essence a recapitulation of the historical misconception that sometimes during the process of evolution species will develop features which literally cause their own extinction.

Puzzle 6c (Diachronic V). This puzzle consists of an illustration of an outcrop, with alternating sets of strata. One set consists of marine-based sedimentary rock that was deposited in the winter and the other set consists of terrestrial-based sedimentary rock, containing fossils of a single dinosaur species, deposited in the summer. A successful solution integrates two cycles: (a) a rock-forming cycle based on sea level changes, and (b) a biological cycle recording the appearance and disappearances of the dinosaurs. Correctly integrated, these cycles suggest that the dinosaurs were migrating to and from this area based on local changes in the environment.

A breakdown of the mean scores received for each of the three diachronic schemes indicates that the limiting factor in this puzzle was the interpretation of cause (changing sealevels) and effect (dinosaur migration); in other words, the scheme of interstage linkage was most critical to this puzzle, as it consistently received the lowest mean scores among the six grades. In contrast, most subjects were able to grasp the cyclical change in depositional environments (a transformation scheme) and subsequently this scheme received the higest mean scores.

Similar to the the results of the two previous fossil sequence puzzles, there were significant differences between Grade 9, 10, 11, and 12 students and Grade 7 and 8 students, respectively (Table 4). However, in addition, Grade 12 students were significantly better than Grade 9 students.

TST Results

The most important finding of this analysis was that there was a strong correlation between Temporal Organization III, which primarily depended on the geological skill of stratigraphic correlation, and the selected spatial visualization puzzles of the MGMP (Table 5). In other words, the abilty to correlate fossiliferous strata temporally is influenced by subjects' ability in spatial visualization. Among the other puzzles no strong correlations were found.

Taken as a whole, there was a slight correlation between the entire suite of GeoTAT puzzles used in this test and the MGMP spatial visualization puzzles ($r = 0.24$). However, each GeoTAT puzzle used in this analysis required a different suite of geological skills and diachronic schemes. For example, the Diachronic II and Diachronic V puzzles differed in their emphasis on different schemes of diachronic thinking, whereas the Temporal Organization I (succession and original horizontality) and Temporal Organization III (superposition and correlation) puzzles focused on different elements of the temporal organization scheme. Thus, low r values in some of these puzzles lowered the overall correlation between the two questionnaires (GeoTAT and MGMP).

SFT Results

As noted previously, results from the absolute time puzzle (Puzzle 4) suggested the possibility that subjects believed that the absolute age of the layers was connected to its proportional size. Observations made in earth science classes as well as in-depth interviews with five students in Grades 9, 11, and 12 strengthened this assumption; to test this conjecture further, the SFT was distributed among a sample of high school earth science majors.

Results from this questionnaire indicate that in general, effects of size both of the entire exposure and the individual layers (Table 6) and number of layers in an exposure (Tables 7 and 8) influenced geology students' understanding of relative age. This is supported by the fact that no

Table 5
Product moment correlation between select puzzles from GeoTAT and MGMP spatial visualization test ($N = 172$)

GeoTAT Puzzle	Geological Skills Required	Correlation (r) with MGMP	p
Temporal Organization I	Superposition, original horizontality	0.09	.24
Diachronic II	Superposition, actualistic thinking, and causal thinking	0.14	.13
Diachronic V	Superposition, actualistic thinking, and causal thinking	0.21	.02
Temporal Organization III	Superposition, correlation	0.41	.0001

Table 6

Distribution of scores for comparison of two outcrops with equal numbers of layers (5) but differences in height (Outcrop B > A)

	Question: Try to Estimate Which of the Two Outcrops Is Older, A or B? ($N = 52$)				
Outcrop chosen	Outcrop A	Outcrop B	No difference	Impossible to know	No answer
Cumulative score	21%	38%	4%	35%	2%
Reasoning	Compaction of layers	Outcrop taller	Same number of layers	Correct	Not applicable

more than 35% of the sample in any of the puzzles correctly answered that it was impossible to estimate the age of an outcrop based on its size or number of layers (Table 6).

Even among subjects who answered correctly, it was rare (8%) for them to explain specifically why it was impossible to estimate the age, i.e., because they lacked critical information, the rate of deposition. Instead, most simply stated that ‘‘data were missing’’ without expressly stating which data were missing, or that this situation depended on rock type or environmental conditions. In the latter case, subjects were partly correct: Environmental conditions do indeed affect depositional rate. Nonetheless, this answer is incomplete in that subjects did not directly state the connection between rate of deposition and environmental conditions.

Amongst those students who incorrectly answered the three puzzles, the following specific alternative frameworks were found.

1. The compaction of rock layers strongly affects their overall age. Thus, in Table 6, when comparing two outcrops with equal numbers of layers but differences in height, about one-third of the incorrect answers directly stated that the smaller outcrop was older. When asked to explain their reasoning, subjects answered that the layers were in a process of compaction.
2. The number of layers even more so than their size, or for that matter the size of the outcrop itself, is of critical importance when determining the relative age of two rock exposures. This is clearly seen in Table 8, in which subjects compared pictures of two outcrops of unequal size in which the smaller outcrop had more layers. Of the total subjects 52% (or 75% of the incorrect responses) incorrectly chose the smaller outcrop with the larger number of layers.

Discussion

In this study we attempted to define the factors which affect a student’s ability to reconstruct geological systems, such as depositional environments and fossil sequences, that have developed

Table 7

Distribution of scores for comparison of two outcrops with unequal numbers of layers (Outcrop A = 10; B = 5) and no differences in height

	Question: Try to Estimate Which of the Two Outcrops Is Older, A or B? ($N = 52$)				
Outcrop chosen	Outcrop A	Outcrop B	No difference	Impossible to know	No answer
Cumulative score	48%	8%	11%	31%	2%
Reasoning	Larger number of layers	Layers are thicker	Same height	Correct	Not applicable

Table 8
Distribution of scores for comparison of two outcrops with unequal numbers of layers (Outcrop A = 10; B = 5) and differences in height (Outcrop B > A)

	Question: Try to Estimate Which of the Two Outcrops Is Older, A or B? ($N = 52$)				
Outcrop chosen	Outcrop A	Outcrop B	No difference	Impossible to know	No answer
Cumulative score	52%	17%		29%	2%
Reasoning	Larger number of layers	Outcrop taller	Size and No. of layers equalizes differences	Correct	Not applicable

over time. The organizing framework for this understanding is the diachronic thinking model of Montagnero (1996), which has been adapted for the specific needs of this study (Figure 1).

In reviewing the results of the full diachronic puzzles from the GeoTAT, we repeatedly see that the transformation scheme is key in that it influences the other two diachronic factors, confirming Montagnero's (1996) observations. To a large degree, this is because, if the subject does not recognize that a transformation in time has taken place, he will not activate the other diachronic schemes. As a concrete example, we can turn to Puzzle 2 (Diachronic II) in which it was difficult to suggest the causal reasons (interstage linkage scheme) which deposited the strata unless one was able to first deduce the local environmental changes based on the transformation scheme which are reflected in the outcrop's composition.

The most important factor in understanding the scheme of transformation was the ability to interpret the clues which permits one to build a connection between the past and the present (Table 1). Indeed, even some of the youngest subjects (Grade 7) understood this basic premise if they were able to grasp the visual and/or literal clues provided in the puzzles. This begs the question as to what the minimum age is for understanding and using actualistic thinking. A definitive answer has practical significance because actualistic thinking is a component of many

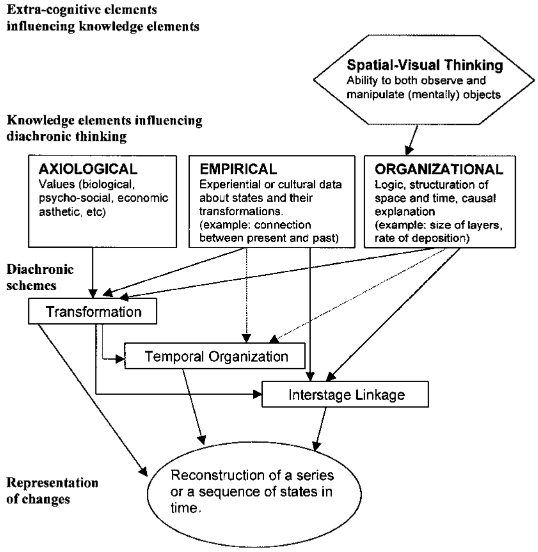


Figure 1. Model of temporal logic in geology (based on Montagnero 1996).

sciences besides the earth sciences. We therefore suggest that more research be invested in this issue.

In contrast to their experience with the scheme of transformation, students had little difficulty in activating the temporal organization scheme when attempting to solve the five full diachronic puzzles. It should be remembered, however, that with the exception of Puzzle 1b (Diachronic I) the strata were undisurbed (i.e., neither folded nor tilted), reducing their complexity. When the strata were folded as in Puzzle 1b, students had much greater difficulty solving the puzzle: so much so, that this puzzle received the lowest marks among the five full diachronic puzzles.

While designing the GeoTAT instrument during the pilot study, we found that both junior and senior high school students had no difficulty in arranging strata when the layering was undisturbed. Indeed Ault (1981) conclusively showed that children as young as Grade 2 possessed the cognitive skills to solve problems involving simple superposition (i.e., in which the strata are undisturbed), so this problem was not tested with this research sample.

This prompts a question as to the roots of this understanding. Ault (1981) argued that the understanding of superposition is based on an understanding of succession, or in other words a series of before and after relationships. He based his argument on Zwart (1976), who argued convincingly that the understanding of before and after is the basis of temporal understanding itself. We noted earlier that Stevenson and Pollitt (1986) as well as Harner (1982) showed that very young children (as young as three years in age) understand the verbal meaning of ‘‘before and after.’’ In his study, Ault (1981, 1982) found that kindergarten children could serially order events in time using an expanded version of the before and after concept, termed transitive thinking, which states that if *a* is before *b* and *b* is before *c*, then *a* is before *c*. The question remains as to whether children transfer this verbal understanding into the visual reality of strata.

In this research, the sample was tested with an example of superposition (Puzzle 1a) in which the strata were complicated by folding. To solve it, students required the use of a second scheme of temporal organization—the principle of original horizontality. In using this principle, the subject ignored his natural inclination that the height of the strata is necessarily equivalent to its relative age.

In the results, the youngest subjects (Grade 7–8) had the most difficulty in recognizing the need for the principle of original horizontality, and thus they were more susceptible to misleading clues of relative height even when given a strong written clue which might overcome this problem. In junior high schools in Israel, students have recently begun studying geology using the curriculum *The Rock Cycle* (albeit not the subjects who took part in this study). In this program the most complicated example of superposition they encounter is tilted rather than folded strata. Although their understanding of superposition has not been tested in a pre/post fashion, their observed work in class indicates that they understand its internal logic and, more important, can apply it in the field (Kali, 2000).

A third level of complexity after the principles of superposition and lateral horizontality in temporal organization was tested in the GeoTAT through stratigraphic correlation (Puzzle 5). Based on their mean scores, the youngest two groups (Grade 7–8) relied primarily on a strategy of either random placement or their primitive understanding of superposition for correlating the fossils. This explains the significant differences with the older (Grade 9–12) subjects.

This result contrasted somewhat with Ault (1981), who claimed that at least some children as young as Grade 2 understood the basic logic of correlation. In interviews with such children, they were able to order the layers of a simulated compost pile using core samples at different depths. Ault’s (1981, 1982) example is difficult to compare with the present study because his problem is less complex and does not represent a true geological situation. Moreover, his sample size was small, a total of 10 children interviewed, in which only 3 were successful in Grade 2. Even among the Grade 6 students, only 5 of 10 were successful in this task. Moreover, he admitted that it was "rare" for such students to be able to correlate strata in the field.

Based on the mean grade scores obtained in this study, we suggest that the limit for stratigraphic correlation is no earlier than Grades 7–8. Nonetheless, with instruction it is possible to improve this skill among junior high school students. In Academic Year 1998–1999, two academically low-achieving junior high school classes participated in implementation of the learning program "From Dinosaurs to Darwin" (Dodick & Orion, 2002a). Comparison of preprogram mean scores (Class 1 = 22.4%; Class 2 = 55.0%) with postprogram mean scores (Class 1 = 51.4%; Class 2 = 75.5%) indicated that the test classes had significantly improved ($p < .05$ level) in their ability to correlate strata.

As noted in the Results, the cognitive dividing line in stratigraphic correlation was in integrating strata from different localities. To do so, the subject must employ a strategy which involves both superposition when warranted and translation across the rows in search of matching fossils. In other words, test subjects needed to employ a much more three-dimensional strategy than with the complex superposition puzzle (Puzzle 1a). This suggests that one factor limiting success on stratigraphic correlation is subjects' ability in visual-spatial perception.

This possibility was investigated with the TST instrument in which results from a select number of puzzles from the GeoTAT were correlated against results obtained with the MGMP, a spatial visualization test. Results from the TST suggest that in the case of stratigraphic correlation a subject's ability in spatial visualization affects his ability in temporally ordering the strata.

To those who argue that such ability is simply an artifact of general intelligence (a criticism of any correlation analysis) and not specifically tied to subjects' ability in spatial visualization, we suggest that the disparity in mean scores among the four puzzles counters this criticism (Table 5). If such ability were indeed tied to general intelligence, the correlation coefficients would have been similar with all four selected puzzles of the GeoTAT. Moreover, the puzzle with the highest correlation coefficient is the puzzle which truly required a three-dimensional strategy (i.e., stratigraphic correlation), incorporating both vertical movement along the outcrops as well as translation between outcrops. In contrast, the superposition puzzle required a one-dimensional strategy of simply moving up the stratigraphic column and hence received the lowest correlation coefficient.

Intermediate r values were found for the two full diachronic puzzles (Diachronic II and Diachronic V) because they required what may be called a two-dimensional approach. In such puzzles, the students had to move up the stratigraphic column based on superposition while integrating the lateral movements of the sea and/or dinosaurs into alternating layers of the stratigraphic column.

In the five full diachronic puzzles of the GeoTAT, temporal organization was based on simple superposition on a single stratigraphic column. Multiplying the number of columns and adding in factors of faulting, folding, and crosscutting relationships in such diachronic puzzles would likely require greater spatial visualization ability. This relationship should therefore be investigated further, with advanced earth science students. (Such questions were deemed to be too complex for the sample tested in this study.)

A second factor that should be investigated is the effect of fieldwork. All of these puzzles were tested with written instruments; in the field it is possible that even an undisturbed (i.e., unfolded and untilted) outcrop might require higher levels of spatial ability than its paper and pencil counterpart. Indeed, we have noted such problems in field exercises with both students and teachers.

Interestingly, this is not the first research to suggest a connection between spatial visualization and temporal understanding. In the 18th century, Kant proposed such a relationship in his *Critique of Pure Reason* (Friedman, 1990). More recently Friedman (1983, 1989, 1992) proposed that time structures with a definite pattern, such as the days of the week or months of the year, are represented in the mind as analogous to positions in space. ‘‘This position like information when accessed can reveal ‘where’ a given element occurs relative to other elements in the pattern’’ (Friedman, 1992, p. 69).

A third temporal organization puzzle (Puzzle 4, Temporal Organization II) testing student understanding of absolute time measurement was tested with both the nongeology and geology majors. As noted previously, the vast majority of the nongeology majors were unable to solve this puzzle; indeed even the geology majors had great difficulty with it.

These results are mentioned because they represent a general problem that should be considered when discussing the significance of strata to geological time. As was noted, almost all of the subjects tested partitioned the strata into equal portions of time, almost as if they were units on a ruler. This suggested that the subjects believed that the absolute ages of the strata were proportional to their size.

To test this assumption, a third questionnaire, the SFT, was distributed among a sample of high school earth science majors. This questionnaire confirmed the hypothesis that even students with a background in geology put undue emphasis on both size and numbers of layers when considering the age of geological layers.

Because in the original puzzle (Puzzle 4 of the GeoTAT) they apportioned equal amounts of time to the strata, it is possible that the subjects believed that geologic deposition occurs at a uniform or linear rate. In other words, they do not understand the concept of rates of change.

In fact, as geologists know full well, sedimentation is not a uniform process; changing environmental conditions can drastically affect the rates of such processes. Moreover, sedimentation is often halted and erosion takes over, leaving large gaps in the rocks’ temporal record. Such an erosional surface separating adjacent rock beds is termed an unconformity by geologists, and it represents time which has been lost from the geological record (Press & Siever, 1982).

Obviously, such changes in sedimentation rate complicate one’s understanding of rock formation; however, it is a critical element students should comprehend if they are to build a complete understanding of this process. Indeed, the understanding of rates of change is a basic concept of all scientific disciplines both as a methodological problem (the measurement of rates) and as a philosophical problem (continuous vs. discrete rates of change). Thus, exposing students to this concept within the earth sciences gives them a better understanding of one of science’s universal concepts.

Unfortunately, there has been no research within science education on the relationship between changing rates and its effect on earth-based phenomena such as deposition. In fact, most of the research within science education is associated with physics and mathematics education (Karplus, Pulos, & Stage, 1983; Thompson, 1991; Thompson & Thompson, 1992). Such studies are interested in students’ understanding of the physical interrelationships among speed, distance, and time, or dynamic systems that are observable in real time. In contrast, the earth sciences add a level of abstraction in that one must visualize the dynamic processes of the past such as sedimentation, using static clues culled from present-day geological structures.

In terms of trends in age, we found that in 6 of the 10 GeoTAT puzzles there was a significant difference ($p < .05$) between Grade 9–12 students and Grade 7–8 students. These include the two puzzles involving actualistic thinking (Puzzles 6a and 6b), the puzzle involving stratigraphic correlation (Puzzle 5), and three puzzles involving the full suite of diachronic thinking (Puzzles 3a, 3b, and 6c).

This suggests that to improve their ability to reconstruct geological systems that have developed over time, programs in earth science at the Grade 7–8 level should focus on problems involving stratigraphic correlation and the basic principles of diachronic thinking. In teaching such problems, close attention should be paid to students' cognitive understanding of the diachronic schemes as well the content knowledge which is the basis for activating each scheme. Thus, if the subject does not recognize that limestone forms in shallow water, he may not be able to reconstruct the stages in which a series of limestone strata formed. With increasing age, it is probable that such knowledge factors take on greater importance, as the child better assimilates the logic of the diachronic schemes.

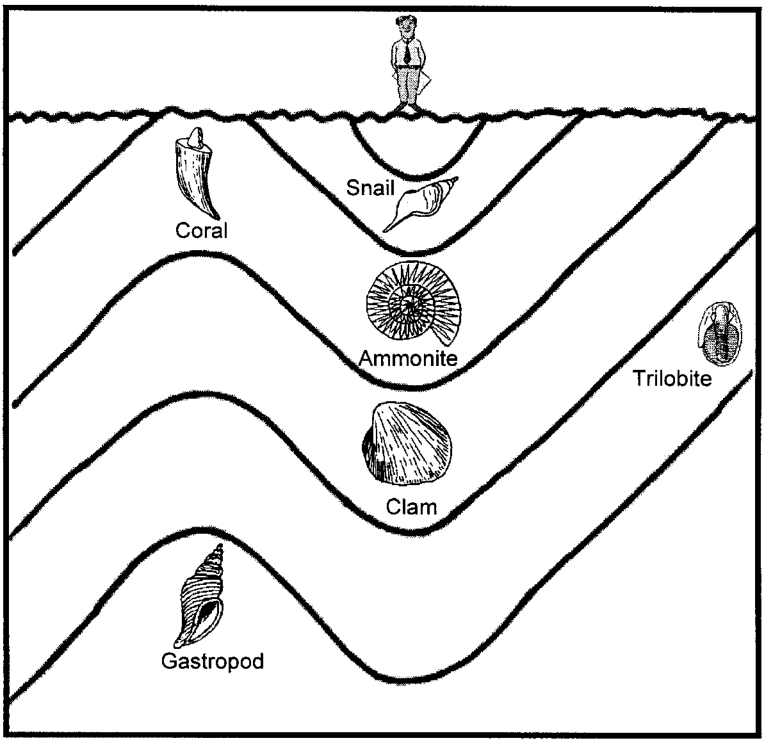
Attention should also be given to those extra cognitive factors which complicate a student's diachronic thinking processes. As was noted previously, strata are three-dimensional structures. Proper attention must therefore be given to improving a subject's understanding of spatial relationships, which in turn will affect his understanding of temporal organization. This factor becomes more important as the geological structures become more complex (folded, faulted, multiple exposures) or when operating on structures in the field, the true test of all geological understanding.

In summary, the model delineated in this study can be used to focus on factors which limit a student's understanding of transformations in time. Thus, it has practical value to both curriculum designers and teachers who are interested in educating students in any science that is contrained by historical factors including geology, palaeontology, ecology, archaeology, and even astronomy.

Appendix

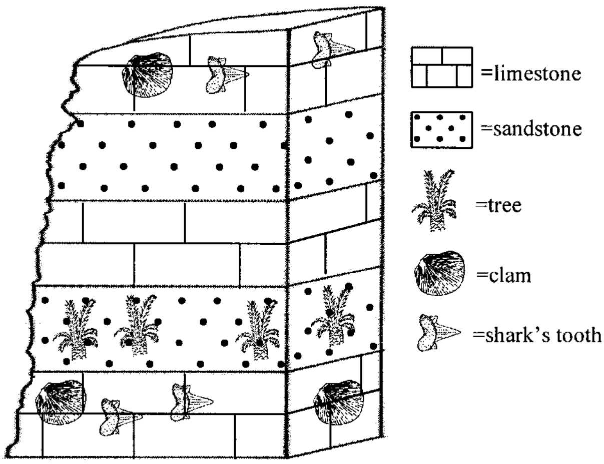
GeoTAT Questionnaire

1. The geologist in the diagram below is standing on a column of marine sedimentary rock containing fossils.

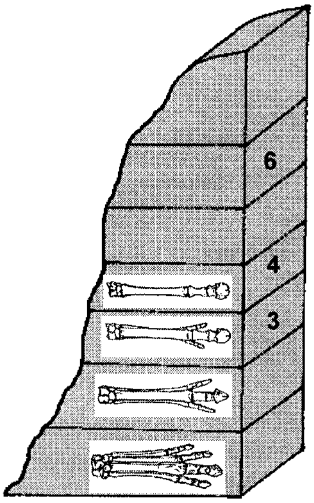


1a. Attempt to order the fossils according to their age from the oldest fossil to the youngest fossil. (Clue: marine sedimentary rock is originally deposited in horizontally lying layers.)
1b. Try to reconstruct the processes in order which lead to the creation of the rock exposure in the picture above.

2. The illustration below represents a series of rock layers from a specific locality in the world. Try to reconstruct the stages which created this sequence of layers based on their order of formation.

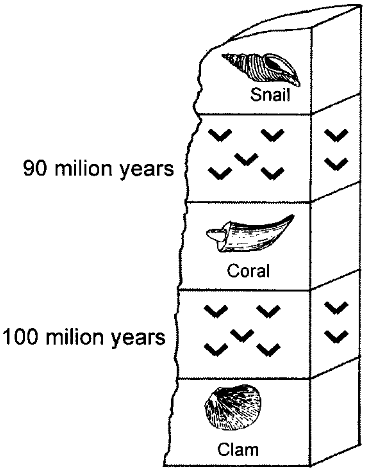


3. The illustration below represents a fossiliferous rock exposure. The fossils are the remains of bones from the feet of unidentified species of mammals.

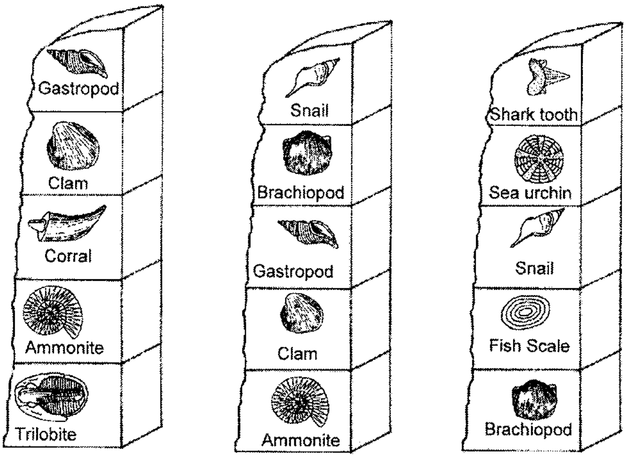


3a. Try and describe the process that took place between Rock Layer 1 to Layer 4.
3b. Try to suggest two possible reasons for the absence of fossils after Rock Layer 4.

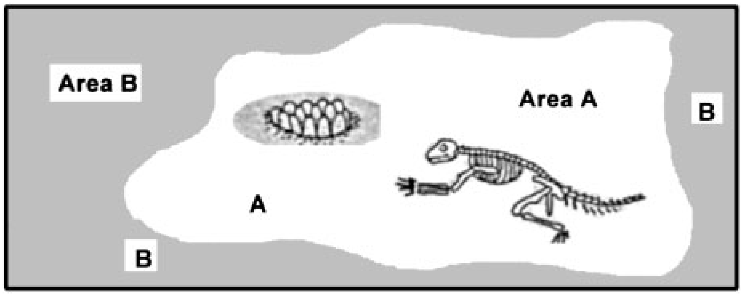
4. The following picture represents a rock exposure that contain three types of fossils (snail, coral, and clam). Two layers of igneous rocks (represented by the symbol "V") lie between the layers containing the fossils. The age of the igneous rock layers have been determined in the lab by scientists. Try to determine the absolute age in years of the three different fossils (snail, coral, and clam).



5. The illustration below represents three rock exposures containing fossils. Try to order the fossils according to their implied age, from the oldest fossil to the youngest fossil.



6. The following illustration represents a dinosaur excavation site. This excavation can be broken down into two areas.

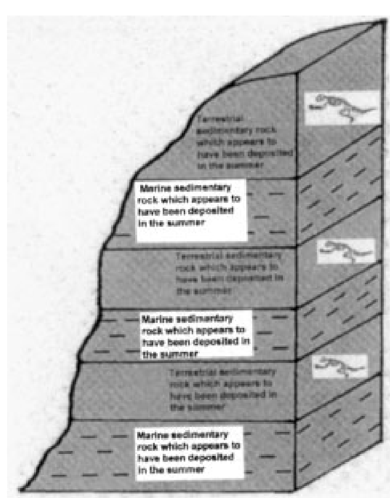


Area A. This site is built of terrestrial sedimentary rock containing the skeletons of dinosaurs. Two important points can be noted about this area.

1. The dinosaur skeletons excavated in this area range in size from very tiny to very large. This suggests that in this one area, the age range of the dinosaurs was broad ranging from newly hatched babies to fully grown adults.
2. In this site, a large series of nests containing fossilized eggs was discovered.

Area B. This area surrounds area B and is built of marine sedimentary rock containing fish.

6a. Try to reconstruct how this looked when the dinosaurs were alive. What did Areas A and B comprise?
6b. What in your opinion might be the significance of the fact that in one single area scientists found a species of dinosaur ranging in size and age from egg to adult?
6c. When scientists excavated this area deeply they found an alternating arrangement of layers consisting of marine sedimentary rock containing no fossils and terrestrial sedimentary rock containing fossils of dinosaurs (in the illustration below). What is the significance of this alternating arrangement of layers containing terrestrial sedimentary rock containing dinosaurs, and marine sedimentary rock without dinosaurs?



References

Ault, C.R. (1982). Time in geological explanations as perceived by elementary school students. Journal of Geological Education, 30, 304–309.

Ault, C.R. (1981). Children's concepts about time no barrier to understanding the geologic past. Unpublished doctoral dissertation, Cornell University, Ithaca.

Ben-Chaim, D., Lappan, G., & Houang, R.T. (1986). Development of a spatial visualizaton 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 on spatial visualization skills of middle school boys and girls. American Educational Research Journal, 25, 51–71.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In Friedman, W. (Ed.), The developmental psychology of time (pp. 209–255). New York: Academic Press.

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

Dodick, J.T. & Orion, N. (2002a). From dinosaurs to Darwin: Evolution from the perspective of "deep time." A curriculum for the Science for All program. Rehovot, Israel: Weizmann Institute of Science.

Dodick, J.T. & Orion, N. (in press). Measuring student understanding of geological time. Science Education.

Fraisse, P. (1982). The adaptation of the child to time. In Friedman, W. (Ed.), The developmental psychology of time (pp. 113–140). New York: Academic Press.

Friedman, W. (1992). The development of children's representations of temporal structure. In Macar, F. (Ed.), Time action and cognition (pp. 67–75). Amsterdam: Kluwer Academic.

Friedman, W. (1990). About time: Inventing the fourth dimension. Cambridge: MIT Press.

Friedman, W. (1989). The representation of temporal structure in children, adolescents, and adults. In Levin, I. & Zakay, D. (Eds.), Time and human cognition: A life span perspective (pp. 259–304). Amsterdam: Holland.

Friedman, W. (1983). Image and verbal processes in reasoning about the months of the year. Journal of Experimental Psychology: Learning Memory and Cognition, 9, 650–666.

Friedman, W. (1982). Conventional time concepts and children's structuring of time. In Friedman, W. (Ed.), The developmental psychology of time. New York: Academic Press.

Friedman, W. (1978). Development of time concepts in children. In Reese, H.W. & Lipsett, L.P. (Eds.). Advances in child development and behaviour, vol. 12 (pp. 267–298). New York: Academic Press.

Frodeman, R.L. (1996). Envisioning the outcrop. Journal of Geoscience Education, 44, 417–427.

Frodeman, R.L. (1995). Geological reasoning: Geology as an interpretive and historical science. Geological Society of America Bulletin, 107, 960–968.

Gould, S.J. (1987). Time's arrow, time's cycle. Cambridge: Harvard University Press.

Happs, J.C. (1982). Some aspects of student understanding of two New Zealand landforms. New Zealand Science Teacher, 32, 4–12.

Harner, L. (1982). Talking about the past and the future. In Friedman, W. (Ed.), The developmental psychology of time (pp. 141–169). New York: Academic Press.

Huck, S.C., Cormier, W.H., & Bounds, W.G. (1974). Reading statistics and research. New York: Harper and Row.

Kali, Y. (2000). The rock cycle. Unpublished doctoral dissertation, Weizmann Institute of Science, Rehovot, Israel.

Kali, Y. & Orion, N. (1996). Spatial abilities of high school students in the perception of geologic structures. Journal of Research in Science Teaching, 33, 369–391.

Karplus, R., Pulos, S., & Stage, E. (1983). Early adolescents' proportional reasoning on "rate" problems. Educational Studies in Mathematics, 14, 219–234.

Marques, L.F. & Thompson, D.B. (1997). Portuguese students' understanding at age 10/11 and 14/15 of the origin and nature of the Earth and the development of life. Research in Science and Technology Education, 15, 29–51.

Mayer, V.J. & Armstrong, R.E. (1990). What every 17 year old should know about planet earth: The report of a conference of educators and geoscientists. Science Education, 74, 155–165.

Montagnero, J. (1996). Understanding changes in time. London: Taylor and Francis.

Montagnero, J. (1992). The development of the diachronic perspective in children. In Macar, F. (Ed.), Time action and cognition (pp. 55–65). Amsterdam: Kluwer Academic.

Montgomery, C. (1993). Physical geology. New York: William. C. Brown.

Noonan-Pulling, L.C. & Good, R.G. (1999, March). Deep time: Middle school students' ideas on the origins of earth and life on earth. Paper presented at the National Association for Research in Science Teaching annual meeting, Boston, MA.

Oversby, J. (1996). Knowledge of earth science and the potential for development. School Science Review, 78, 91–97.

Piaget, J. (1969). The child's conception of time. London: Routledge and Kegan Paul.

Press, F. & Siever, R. (1982). Earth. New York: W.H. Freeman.

Roseman, J. (1992). Project 2061 and evolution. In Good, R., Trowbridge, J.E., Demastes, S., Wandersee, J., Hafner, M., & Cummins, C. (Eds.), Proceedings of the 1992 Evolution Education research conference (pp. 214–229). Baton Rouge: Louisiana State University.

Rowland, S.M. (1983). Fingernail growth and time–distance rates in geology. Journal of Geological Education, 31, 176–178.

Ritger, S.D. & Cummins, R.H. (1991). Using student created metaphors to comprehend geological time. Journal of Geological Education, 39, 9–11.

Schoon, K.J. (1989, March/April). Misconceptions in the earth sciences: A cross-age study. Paper presented at the 62nd Annual Meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Stevenson, R.J. & Pollitt, C. (1986). The acquisition of temporal terms. Journal of Child Language, 14, 533–545.

Thompson, P.W. & Thompson, A.G. (1992, April). Images of rate. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Thompson, P.W. (1991). The development of the concept of speed and its relationship to concepts of rate. In Harel, G. & Confrey, J. (Eds.), The development of multiplicative reasoning in the learning of mathematics. Albany: SUNY Press.

Trend, R. (2000). Conceptions of geological time among primary teacher trainees, with reference to their engagement with geosciences, history and science. International Journal of Science Education, 22, 539–555.

Trend, R. (2001). Deep Time Framework: A preliminary study of UK primary teachers' conceptions of geological time and perceptions of geoscience. Journal of Research in Science Teaching, 38, 191–221.

Trend, R. (1998). An investigation into understanding of geological time among 10- and 11-year old children. International Journal of Science Education, 20, 973–988.

Trend, R. (1997, July). An investigation into understanding of geological time among 10- and 11-year old children, with a discussion of implications for learning of other geological concepts. Paper presented at Second International Conference on Geoscience Education and Training, Hilo, HI.

Zwart, P.J. (1976). About time: A philisophical inquiry into the origin and nature of time. Amsterdam: North-Holland.