Measuring Student Understanding of Geological Time

JEFF DODICK, NIR ORION
Department of Science Teaching, Weizmann Institute of Science, Rehovot 76100, Israel

Received 10 January 2001; revised 11 May 2002; accepted 25 June 2002

ABSTRACT: There have been few discoveries in geology more important than "deep time"—the understanding that the universe has existed for countless millennia, such that man's existence is confined to the last milliseconds of the metaphorical geological clock. The influence of deep time is felt in a variety of sciences including geology, cosmology, and evolutionary biology. Thus, any student that wants to master these subjects must have a good understanding of geological time. Despite its critical importance, there has been very little attention given to geological time by science education researchers. Of the work that has been done, much of it ignores the cognitive basis for students' understanding of geological time. This work addresses this gap by presenting a validation study for a new instrument—the GeoTAT (Geological Time Aptitude Test). Consisting of a series of open puzzles, the GeoTAT tested the subjects' ability to reconstruct and represent the transformation in time of a series of geological structures. Montagnero (1992, 1996) terms this ability "diachronic thinking." This instrument was distributed to a population of 285 junior and senior high school students with no background in geology, as well as 58 high school students majoring in geology. A comparison of the high school (grades 11–12) geology and non-geology majors indicated that the former group held a significant advantage over the latter in solving problems involving diachronic thinking. This relationship was especially strengthened by the second year of geological study (grade 12), with the key factor in this improvement being exposure to fieldwork. Fieldwork both improved the subjects' ability in understanding the 3-D factors influencing temporal organization, as well as providing them with experience in learning about the types of evidence that are critical in reconstructing a transformational sequence. *Sci Ed* **87:**708–731, 2003; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/sce.1057

INTRODUCTION

In the history of geology there have been two discoveries, plate tectonics and "deep time," which have literally shaped the core of this discipline. By "deep time" it is meant the understanding that the universe has existed for countless millennia, such that man's earthly dominion is confined to the last milliseconds of the metaphorical geological clock (Dodick & Orion, in press; Gould, 1987).

The influence of deep time is felt in a variety of scientific disciplines including geology, cosmology, and evolutionary biology. Thus, any scientist or student that wants to master any of these subjects must first have a good understanding of "deep time."

Correspondence to: Jeff Dodick; School of Education and Social Policy, Northwestern University, 2120 Campus Dr., Evanston, IL 60208, USA; e-mail: j-dodick@northwestern.edu

Despite the critical importance of this concept, there has been very little attention given to it by science education researchers. The small amount of research that has been completed can be roughly divided into two groups: event-based studies and logic-based studies.

In the category of event-based studies, the researchers include all research that surveys student understanding of the entirety of deep time (that is beginning with the formation of the earth or the universe). In such studies, the general task is sequencing a series of events (for example the first appearance of life on earth) absolutely, along a time-line, or relatively using picture-sorting tasks. Often in such sequencing tasks, the subject is asked to justify his reasons for his proposed temporal order. Such studies include Noonan-Pulling and Good's research (Noonan-Pulling & Good, 1999) on the understanding of the origins of earth and life amongst junior high school students; a similar study by Marques and Thompson (1997) with Portuguese students; and Trends' studies respectively on the conception of geological time amongst 10–11-year-old children (Trend, 1997, 1998) as well as amongst primary teacher trainees (Trend, 2000).

The difficulty with such studies is that they largely reflect the subjects' knowledge of particular events rather than the underlying cognitive basis for their understanding of geological time itself. This second approach is found in the second designated type of research, the logic-based study. Unfortunately, the pioneering study of Aults (1981, 1982) remains the only in-depth research of this type. In this work, he interviewed a group of students (grades K–6) using a series of puzzles that tested how they understood (and could reconstruct) a series of geological strata. On the basis of Zwart's suggestion that the development of man's temporal understanding lies in the before and after relationship (Zwart, 1976), Ault (1981) theorized that children organize geological time relationally.

Based on his findings, Ault (1981, 1982) claimed that young (grade 2–6) children's concept of conventional time was no impediment towards their 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 field, these same children had difficulties in solving similar types of problems, indicating that there was little transfer from classroom problems to authentic geological settings.

These difficulties can be traced to Ault's research design (Ault, 1981), which was influenced by Piaget's work on time cognition (Piaget, 1969). According to Piaget (1969), a young child's understanding of time is tightly bound to his concept of motion; thus, the research problems he used were taken from physics. However, the geological sciences builds its knowledge of time through visual interpretation of *static* entities (such as formations and fossils) (Frodeman, 1995, 1996). Ault's design, involving motion, simply multiplied the variables that he needed to explain (Ault, 1981), as he admitted in a later work (Ault, 1982). Further, it did not concentrate its efforts on the special qualities of geological time (most importantly its enormous scale) that might complicate a young child's thinking.

Indeed, there is no reason to suggest that an understanding of the (logical) relationships amongst strata should necessarily allow one to both conceptualize and internalize the entirety of geological time. In contrast, the authors will argue that the two forms of understanding can be studied as separate entities.

In addition to the studies noted above, one might add the small body of research which catalogues general misconceptions in geology, including research problems related to geological time (Happs, 1982; Marques, 1988; Oversby, 1996; Schoon, 1989). The obvious problem with such studies is that they do not cognitively model student understanding of geological time.

Finally, one might mention those works within geological education which have concentrated on the practical elements of teaching the (geological) scale of time (Everitt, Good,

& Pankiewicz, 1996; Hume, 1978; Metzger, 1992; Ritger & Cummins, 1991; Rowland, 1983; Spencer-Cervato & Day, 2000). Unfortunately, these teaching models have never been critically evaluated, so they are of untested value to the pedagogic literature. It might be noted, however, that one model (Ritger & Cummins, 1991) does show promise in that it emphasizes a student-centerd (constructivistic) approach in which the student builds a "personal metaphor" of geological time which allows him to structure this abstract concept based on his own criteria.

Amongst the studies mentioned above, most involved qualitative research (structured interviews) with small populations of younger students (either in elementary or junior high school). A survey of the science education literature indicates that there has never been a large-scale quantitative study of older student's (junior high to senior high) understanding of geological time. In Israel, students encounter the earth sciences at the junior and senior high school levels; thus, it was important to design an effective, reliable, and valid tool that might be used to check student understanding of a central feature of the earth sciences.

PURPOSE OF THE STUDY

The purpose of this study was to develop a reliable, valid, and sensitive instrument for testing student understanding of geological time (designated GeoTAT for Geological Time Assessment Test). The GeoTAT instrument was intended for the following purposes:

1. Testing the contribution of learning geology towards the development of temporal thinking. (On a practical level such testing permits one to make recommendations towards improving earth-science curricula).
2. Testing the general cognitive abilities of non-geology majors, both in junior and senior high schools, in understanding geological time. Thus, this instrument was also designed to identify the factors that influence the subjects logical understanding of geological time. Moreover, this instrument should be sensitive to differences in age and sex.

THE DEVELOPMENT OF THE QUESTIONNAIRE

Orion and Hofstein (1991), on the basis of the earlier work of Gardner (1975) and Koballa (1984), have delineated a series of stages in designing a reliable and valid tool for measuring student attitudes. In this study, the researchers were interested in cognitive understanding, rather than attitudes; nonetheless, the development scenario of Orion and Hofstein (1991) is general enough so that it may be applied to a wide variety of educational questionnaires. On the basis of their design, the researchers in this study delineated the following stages in developing the GeoTAT:

1. Conceptualization
2. Item formulation
3. Content validation
4. Statistical analysis, which can be divided into two groupings:

    a. Construct validity (including factor analytic investigation and Cronbach's α reliability coefficient).
    b. Sensitivity towards factors of discipline (i.e. learning in the earth sciences), age (grade), and sex.

5. Comparison of an expert's judgement with the statistical analysis.

Conceptualization

The first stage in the development of this instrument was to identify the major factors that might influence student understanding of geological time. This proceeded in two stages:

1. Identifying the general types of geological thinking used in the earth sciences to understand temporal change. This thinking is outlined in the section titled **Geological Time: A Framework for Understanding**.
2. Translating such geological thinking into general cognitive principles. This step was based on the assumption that temporal understanding in geology has a basis in more generalized cognitive principles. Moreover, since students lack a professional understanding of the earth sciences, it is likely that they would rely on such general cognitive techniques in solving temporal problems related to geology. This permits the researcher to better identify students' problems in both logic and lack of knowledge. This process is discussed in the section titled **Cognitive Thinking and Geological Time**.

Geological Time: A Framework for Understanding. As we noted previously, geological time is a cornerstone of the earth sciences. Indeed every investigation that professional geoscientists do is impacted upon by the background of geological time. Yet how do earth scientists make sense of geological time? This section serves as a short introduction into the methods that are used in order to puzzle out the history of the earth.

In most standard textbooks of geology, reconstructing earth history is divided into two areas of focus: relative dating and numerical dating. We will treat each of these topics in turn.

Relative dating pertains to qualitative age relationships in which a geological structure (such as a sedimentary rock layer) is determined to be older, equal in age, or younger than another feature. Relative ages are established through the application of a small number of universal principles. The two most important and best known of these principles were first elucidated by the seventeenth century Danish natural philosopher Nicholas Steno (1638–1686) (Press and Siever, 1998, p. 219):

1. The *principle of superposition* states that each layer of a sedimentary rock in a tectonically undisturbed sequence is younger than the one below it and older than the one above it.
2. The *principle of original horizontality* states that sediments are deposited as essentially horizontal beds. Thus, if we find a sequence of sedimentary rock that is folded or tilted we know that these rocks were deformed after the sediments were deposited.

These two principles allow a geologist to view a vertical set of layers (or stratigraphic sequence) as a chronological record of the geological history of a single region. However, such stratigraphic principles cannot be used to determine the relative ages of geological layers that are widely displaced geographically. This is due to the fact that the same rock units are not continuous around the world. To overcome this problem, geologists turn to the fossil record.

In 1799 (and possibly earlier), the English geological surveyor William Smith established the principle that individual geological strata could be identified by their characteristic fossil contents (Oldroyd, 1996). This is due to the fact that different species of plants and animals evolve over time; preserved within sediments, such organisms create a stratigraphic ordering of fossils known as a *faunal succession* (Press & Siever, 1998).

Smith was the first person to use faunal succession to *correlate* (or match) rocks from different outcrops. In each outcrop, he identified distinct formations—a series of rock layers that almost everywhere have the same physical properties and that contains the same assemblages of fossils. Such a distinctive set of rock layers can be recognized and mapped as a unit.

Using his knowledge of faunal succession, Smith matched up the similarly aged formations found in different outcrops. He was able to do so because such outcrops contained similar assemblages of fossils. By noting the vertical ordering in which the formations were found in each place, he compiled a composite stratigraphic sequence for an entire region. Using this approach, geologists of the last two centuries have been able to correlate strata around the world so as to create a relative geological time scale.

In contrast to relative dating, numerical dating refers to the age measured in years of a geological event or feature. Numerical dating is dependent on the fact that certain types of rock (usually igneous) incorporate within their crystal matrix radioactive materials (such as uranium).

Determination of geologic age is made possible by the fact that radioactive materials decay at a uniform rate, and thus form a geologic clock. This uniform decay rate is commonly stated in terms of a radioactive element's half-life, the time required for one-half of the original number of radioactive atoms to decay. Thus, at the end of the first half-life after a radioactive element is incorporated into a new mineral, half the number of parent atoms remain (replaced by the newly formed daughter isotope atoms). If we know the decay rate, and we can count the number of daughter atoms, as well as the remaining parent atoms, then we can calculate the time that has elapsed since the radioactive clock began to tick (Harbaugh, 1968). In effect we can work back to the time of the formation (and thus age) of the rock material.

As we noted above, radioactive materials for the most part are incorporated into igneous rocks, whereas fossils are found for the most part in sedimentary rocks. This leads to the question of how geologists provide ages for sedimentary rocks and the fossils that they contain? The most common method (and the method tested with the subjects in this research) is through bracketing. Bracketing depends on the fact that layers of sedimentary rocks are sometimes bracketed between layers of dateable igneous rocks. Thus, if the lower igneous layer is dated to 100 million years and the upper igneous layer is dated to 90 million years in age, this means that the sedimentary layers and the fossils they contain are somewhere between 100 and 90 million years in age.

Relative and numerical dating principles permit geoscientists to temporally order strata. However, more than just ordering events, geology is at its very core an historical science, interested in reconstructing the processes that have affected the earth's environment and biota. To do so, geologists had to develop a conceptual tool, termed *uniformitarianism*, which permitted them to reconstruct dynamic processes based on the static evidence of rocks and fossils. Uniformitarianism is a complex concept, which carries a number of different definitions depending on who is using it. For the purposes of this work, it will be defined in the following way: "the geologic processes we see in operation, as they modify the earth's crust today have worked much the same way over geologic time" (Press & Siever, 1998, p. 4). This means that for the most part, geologic processes of the past might be modeled on presently acting processes.

The first use of this principle is seen in the work of the eighteenth century Italian scholar Arduino (1735–1795); however, it was more fully explored in the work of the Scottish geologists James Hutton (1726–1797) and especially Charles Lyell (1797–1875). [Lyell made uniformitarianism a centerpiece of his major opus *The Principles of Geology* (1830–1833), although it was the English philospher William Whewell (1794–1866) who coined the term (Whewell, 1837).]

Aside from exploring the key methodological principle in geology, both Hutton and Lyell gave great thought to the nature of geologic time. In the age of Hutton the standard belief was that the earth was no more than 6000 years in age on the basis of research in biblical chronology. However, both Hutton and latter Lyell rejected this chronology, providing numerous geological features that bespoke of an immensely old earth.

For this paper, one famous example should suffice. As part of his research for *The Principles*, Lyell made an extensive study of the great volcano Etna in Sicily. Lyell calculated that the most recent volcanic cones on Mount Etna (in Italy) were no less than 12,000 years in age (Lyell, 1830–1833, Vol. 3, p. 97). On the basis of these calculations, it was easy to understand that all of Mt Etna was several hundred thousand years in age (Oldroyd, 1996). If this was the age of a single mountain range, then it was easy to understand that the earth as a whole was incalculably old.

It might be noted that such calculations were not only important to Lyell but to his contemporary Charles Darwin. "Deep time" was a key element of the theory of natural selection, because slow, gradual biological change, which might transform the earth's biota, could only operate through the immensity of geological time.

With better numerical dating tools, as well as greater knowledge of the earth's stratigraphic composition it is now possible to date the earth to approximately 4.5 billion years in age. Although, no longer the vast immeasurable entity of which Hutton spoke, still its vast size makes it difficult for even the best earth scientists to comprehend it, even with scale models. All the more so, we should expect problems when we ask (novice) students to both comprehend its expanse and the events that take place within it.

Cognitive Thinking and Geological Time. After defining the basic principles of temporal thinking in geology, it was possible to translate such principles into a general cognitive model that might be tested in a systematic fashion. This model can be broken down into two broad concepts:

1. A (passive) temporal framework in which large-scale geological and biotic events occur—the geological time scale. For geologists this is built through the understanding of an immense time scale, as well as the events (biotic and geological) it represents. In the cognitive literature this is comparable to Friedman's associative networks (Friedman, 1982, 1990), a system of temporal processing that is 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 and/or experience of such events.
2. An (active) logical understanding of "deep time" which is used to reconstruct past environments and organisms. In the earth sciences, such understanding is mediated by logical principles that temporally organize events, such as superposition, as well as by principles that unlock the environmental change in a given site. Unlike the first concept (noted above), which emphasizes a sequence of events in absolute time, the second concept emphasizes relative time, in other words the logical relationship between different strata and the transformations that they represent. In the cognitive literature this is comparable to Montagnero's model of "diachronic thinking" (Montagnero, 1992, 1996). Montagnero (1992, 1996) defines "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 eventual death and decay) of a tree. However, it is argued here that it might equally be applicable to the deposition (through time) of geological features.

Montagnero (1996) has built a model that defines the structural and functional entities that are activated when diachronic thinking is employed. These consist of four schemes that 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 (see Table 1). The diachronic schemes are presented briefly emphasizing the key elements and their connection to geological logic. [For a full treatment of these schemes see Montagnero (1996).]

It is suggested that in any depositional sequence requiring reconstruction, the entire suite of diachronic schemes will be activated. However, sometimes geologists are required to activate individual diachronic schemes, the most noted examples being when they temporally order a folded or faulted sequence of strata (using superposition and original horizontality), or correlate a series of layers by means of fossil content. In this case, the geologist is

TABLE 1
Diachronic Schemes and Their Geological Correlates

Diachronic Scheme and Its Explanation (based on Montagnero, 1996)	Geological Correlate
Transformation: This scheme defines a principle of change, whether qualitative or quantitative. Quantitative transfo-rmation 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 uniformitarianism ("the present as key to past"), in which geological or biological change is reconstructed through comparison with contemporary biological 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 3-D relationship amongst strata, are used as a means of determining temporal organization.
Interstate linkage: The connections between the successive stages of evolutive phenomena. Such connections are built in one of two ways: 1. Relations between a necessary prerequisite and its sequel 2. Cause and effect relationships	In geology such stages are reconstructed through the combination of uniformitarianism (as defined above) as well as through the use of (scientific) causal reasoning.
Dynamic synthesis: Forming a whole from a set of successive stages which are thus conceived of as a manifestation 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 evaluation.

only activating the scheme of temporal organization. Thus, in building the GeoTAT the researchers attempted to build both puzzles that isolated individual schemes as well as puzzles that were dependent on all of the diachronic schemes. In this way, it was possible to determine which schemes limited the students' ability to understand temporal transformation.

In essence, Montagnero's diachronic model (Montagnero, 1992, 1996) is similar to psychological models of causal reasoning (see for example Bullock, 1985; Bullock, Gelman, & Baillargeon, 1982) on their discussion of the three general principles governing causal reasoning). However, in causal reasoning it generally suffices to link two successive states, whereas diachronic reasoning involves imagining a series of transformations over time. Thus, this model is more appropriate to geological thinking in which transformation often extends over a long series of strata.

According to Montagnero (1992, 1996), children as young as 10–11 years of age are able to activate a full diachronic scheme (based on the four schemes outlined above). The limiting factor to this ability is the subject's knowledge of a given phenomenon. On the basis of structured interviews with children between the ages of 7 and 11, Montagnero (1996) notes 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*: This includes an understanding of dimensions (numbers, space, and time) as well as causal relations.
3. *Axiological knowledge*: Knowledge of transformations based on the subject's own system of values. Thus for young children, systems are often judged as progressing towards or away from a preferred state (for example a tree that is slowly decaying as it grows older).

In terms of geological logic it is suggested that the first source—empirical knowledge—is of primary importance to any understanding of the transformation sequence. Thus, for example, if the subject does not know 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, the researcher attempted to limit the types of knowledge, which might prevent a subject from being able to portray geological transformation. This was done by providing the information directly, or by using information that (supposedly) all of the children understood (for example clams live in the sea, implying a former marine depositional environment).

Question Formulation

Prior to drafting the final version of this questionnaire, an inventory of 16 puzzles was tested amongst a select group of students to determine overall levels of clarity, difficulty, and response time. The different sets of puzzles were tested on a population of one hundred sixteen grade 10 non-geology majors, as well as twenty-one grade 10–11 students majoring in geology. These results were analyzed with the help of three earth science education experts, enabling the authors to reduce the puzzle inventory to the seven multipart cognitive puzzles used in the present version of the GeoTAT.

The puzzles in the GeoTAT were further broken down into the following groupings:

1. *Full diachronic schemes*: Such puzzles (1b, 2, 3a, 6c, and 7 in Table 2) contained all of the diachronic factors mentioned in Table 1 above. They were evaluated on a scale based on the presence or absence of these factors noted. Note that the final diachronic puzzle (7) in this group was not connected to the field of geology, but was instead a

TABLE 2
Geological Skills and Their Corresponding Cognitive Skills Tested in the GeoTAT

Puzzle	Geological Skill(s)	Cognitive Skill(s) (Diachronic Schemes)
6a	Actualistic thinking	Isolated transformation
6b	Actualistic thinking	
1a	Superposition, lateral horizontality	Isolated temporal organization
5	Superposition, correlation	
3b	Causal thinking	Isolated interstage linkage
1b	Superposition, lateral horizontality, actualistic thinking, and causal thinking	Full diachronic schemes (transformation, temporal organization, and intestage linkage)
2	Superposition, actualistic thinking, and causal thinking	
3a	Superposition, actualistic thinking, and causal thinking	
3b	Superposition, actualistic thinking, and causal thinking	
6c	Superposition, actualistic thinking, and causal thinking	
7a–e	(Detective puzzle) not applicable	Full diachronic schemes

"detective" puzzle. This was added as a nonbiased indicator of whether geological training indeed impacts on temporal understanding.

2. *Isolated diachronic schemes*: Such puzzles emphasized isolated skills used in diachronic thinking. In general, the level of difficulty in such isolated puzzles was lower than in those puzzles requiring the full suite of diachronic schemes. The best examples of such puzzles are those testing advanced understanding of temporal organization including superposition (1a) or correlation (5) (Appendix for a copy of this instrument).

Content Validation

The content of the questionnaire was validated in two stages. In the first stage, the questionnaire was reviewed by four earth science education experts who analyzed both the quality of the questions as well how closely each question fit the specific parameters established by the researchers. Following this stage, those questions that required the full diachronic schemes were evaluated by Montagnero to determine whether they indeed fit the model of diachronic thinking (personal communication with Professor Jacques Montagnero).

On the basis of the concept validation, it was possible to classify the geological skills represented in each of the GeoTAT puzzles according to the general cognitive skills they required. This permitted the researchers to compare the content validation (provided by the surveyed experts) with the statistical analysis of the construct validation phase.

Table 2 represents a full classification of the geological skills and corresponding cognitive skills tested in each of the GeoTAT puzzles, as determined in the content validation stage. In this table "geological skill" refers to the principles, and techniques that a geologist (or advanced geological student) would use to temporally reconstruct geological structures and systems. "Cognitive skill" refers to the diachronic schemes that are correlated with these geological principles. Note again that puzzle 7 was not connected to the field of geology, but was instead a "detective" puzzle. It was added as a nonbiased indicator of whether geological training indeed impacts on temporal understanding.

Construct Validation

The "final" version of the GeoTAT was administered at the midpoint of academic years 1998–1999 and 1999–2000 to a population of two hundred eighty-five grade 7–12 students. This population was composed of forty grade 7 students, forty-seven grade 8 students, fifty-nine grade 9 students, forty-one grade 10 students, forty-four grade 11 students, and fifty-four grade 12 students. The high school students were not specifically studying geology as part of their matriculation studies (although they may have been exposed to certain concepts in their geography studies).

Amongst the junior high school populations, as well as the grade 10 students, heterogeneity was insured by the fact that all of these students studied the same curriculum. However, in grades 11 and 12 in the Israeli high school system, students choose their matriculation program. Thus, the researcher made an attempt to test only general "education" (homeroom) classes that might contain a cross section of both science and non-science students.

The puzzles were analyzed using a principal factor analysis with varimax rotation. Three factors were retained covering 35% of the total variance. Loadings exceeding 0.4 were considered when identifying the factors (Table 3).

Cronbach α measurements for each factor as well as that of the entire questionnaire (0.78) indicates that the questionnaire was internally consistent. Both the factor analysis and its comparison with the GeoTAT puzzles are displayed in Table 3.

On the basis of the factor analysis, it is possible to say that the factors obtained by this procedure were similar to those obtained by the content validation. To review:

1. Factor 1 stresses the influence of diachronic thinking as four of the five questions that were designed to test this skill are indeed represented in this grouping. The exception

TABLE 3
Factor Analysis (Varimax Rotation)-GeoTAT Scores for Non-Geology SubGroup, Grades 7–12 ($N = 285$).

Puzzle[a]	Factor 1	Factor 2	Factor 3	Cognitive Skills (Diachronic Schemes)
6a	0.45		0.46	Isolated transformation
6b			0.45	
1a		0.43		Isolated temporal organization
1b		0.42		
5		0.42		
2	0.41			Full diachronic schemes
3a	0.53	0.41		
3b	0.52			
6c	0.59			
7a–e (Detective)	0.55			
Variance	17.4%	10.4%	7.7%	
Cronbach's α	0.78	0.68	0.65	

[a]Note that puzzle no. 4 was not tested with the non-geology subgroup in the factor amalysis and is therefore not included.

was question 1b. Although it was designed to test the full suite of diachronic schemes, because of its complexity most of the non-geology subjects treated it as a problem of isolated temporal organization (based on superposition).

2. Factor 2 stresses the influence of temporal organization (based on the geological skills of superposition and correlation). Note that in puzzle 3a student understanding was strongly influenced by this factor as well. Puzzle 3a required the students to interpret a fossil (evolutionary) sequence. Temporal organization of the strata in this puzzle was determined through the principle of superposition. However, subjects unfamiliar with evolution sometimes organized the strata backwards (from top to bottom) because they believed that the fossil content evolves, towards what they believe was "greater complexity" (i.e. from one toe bone to many toe bones). They thus ignored the correct evidence-stratigraphic orientation.
3. The third factor is dominated by the two puzzles, which were designed so as to test the students' use of actualistic thinking (which correlates most closely to the transformation scheme of "diachronic thinking").

A fourth factor representing causal thinking, as suggested in question 3b, was not found because the students interpreted this puzzle as a continuation of question 3a. In other words they saw it as a continuation of the transformational sequence (requiring the full suite of diachronic skills used in 3a). Indeed an incorrect interpretation in 3a directly influenced an incorrect response in 3b.

TESTING THE GeoTAT FOR ITS SENSITIVITY TOWARDS LEARNING IN THE EARTH SCIENCES

To test the sensitivity of the GeoTAT, this questionnaire was distributed amongst two different populations of students:

Fifty-eight grade 11–12 students majoring in geology as part of their Israeli high school matriculation. They are designated as GS (geology students) in this study. Along with their classroom experience in geology, such students participate in a series of fieldwork exercises, in which among other things, they learn about the change in depositional environments over time.

Ninety-eight grade 11–12 students randomly sampled from "homeroom" classes in three Israeli high schools in urban settings. They are designated as NGS (non-geology students). One class each of the grade 11 ($N = 19$) and grade 12 ($N = 17$) NGS were biology majors. Moreover, many of the remaining NGS were also taking biology as part of their studies. This contrasts with the GS, none of whom were studying biology. This difference is noted because the GeoTAT contains four puzzles that have a direct connection to biology, providing a possible advantage to the NGS.

These two populations were compared by using a two-way ANOVA evaluating the effect of grade (11–12) as well as (disciplinary) background (geology vs. non-geology).

The GeoTAT was administered to all of the test groups towards the midpoint of the second semester. The first author was present while all of the subjects completed the questionnaire so as to help with any difficulties that the subjects might have had. All of the subjects were given 45 min (i.e. one regular classroom period) to complete the GeoTAT.

RESULTS AND DISCUSSION

For the purposes of this analysis, the results were divided according to the specific temporal skills as suggested by the factor analysis.

Puzzles Which Tested the Isolated Scheme of Temporal Organization

The results for two of the temporal organization puzzles (1a and 5 represented in Figures 1 and 2, respectively) were quite similar. In puzzle 1a, which tested superposition, there were significant differences between both grade ($F = 4.33$, df = 152, $p < 0.05$) and discipline ($F = 29.76$, df = 152, $p < 0.0001$). In puzzle 5, which focused on correlation, there were also significant differences between grade ($F = 3.80$, df = 142, $p < 0.05$) and discipline ($F = 5.34$, df = 142, $p < 0.05$).

In addition, there appears to be an interaction effect in puzzle 1a ($F = 18.94$, df = 152, $p < 0.0001$). Post hoc testing indicates that there is a significant difference between the grade 12 GS and NGS in both understanding of superposition ($t = 6.98$, df = 152, $p < 0.01$) tested in puzzle 1a and correlation tested in puzzle 5 ($t = 2.86$, df = 142, $p < 0.01$).

This suggests that the geology students do not fully assimilate the concept of superposition until the second year of geology. This may reflect the cumulative effect of fieldwork participation, which helps to transform the abstract 2-D concept as encountered in texts into the 3-D reality of the field.

In the case of correlation (Figure 2), the geology students do not encounter this concept until grade 12; however, once learned it appears to be mastered. Indeed, students in junior high can generally correlate layers given the proper training. In the program "From Dinosaurs to Darwin" (Dodick & Orion, 2000), two classes of grade 9 students were tested both prior to and after completing the program, with the ability to correlate being significantly better in both classes. Ault (1981) himself even suggests that children as young as grade 2 understand the logical principles of correlation, although such knowledge could not be demonstrated in the practical environment of the field.

Puzzle 1b (Figure 3), the third of the identified temporal organization puzzles, is based on interpreting a series of folded, marine exposures. The results show a significant difference between the GS and NGS ($F = 15.43$, df = 99, $p < 0.001$). At the same time there was no significant difference between the grades, nor was any interaction detected.

The source of the NGS difficulties was that they treated this question as a simple case of superposition, ignoring for the most part the folded and eroded nature of the layers. Indeed this is the reason that based on the factor analysis this puzzle is grouped within the second factor (Table 3) as an example of isolated temporal organization, rather than as a full diachronic puzzle as it was originally intended. In structure it is much more closely allied to GeoTAT puzzle 2, which involves the interpretation of depositional environments.

In fact, puzzle 1b might be broken down into a series of transformations (deposition, folding, and erosion) all of which contain the full four schemes of diachronic thinking. This multiplied the factors which the subjects needed to consider, and thus this was the

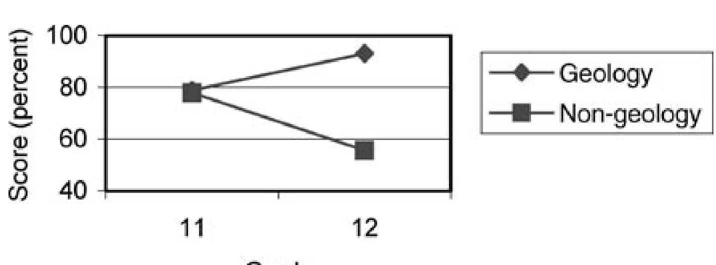


Figure 1. Scores for the GS and NGS on GeoTAT puzzle 1a (superposition).

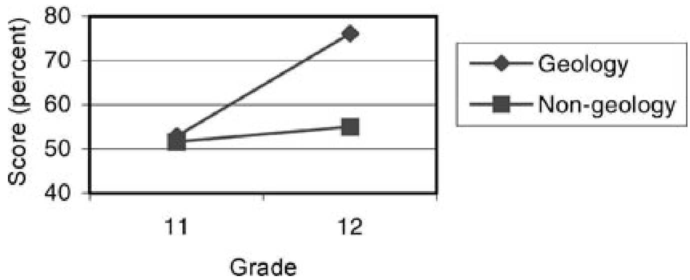


Figure 2. Scores for the GS and NGS on GeoTAT puzzle 5 (correlation).

most difficult puzzle for the NGS to solve. In contrast, the GS easily recognized the factors involved in this puzzle and were therefore significantly better at solving it.

Puzzles Which Tested the Full Suite of Diachronic Schemes

In general, the scores in the five geology-based diachronic puzzles were lower than the other geology-based puzzles in the GeoTAT, for the most part because they encompassed the entire suite of diachronic schemes.

To facilitate further discussion of the diachronic thinking puzzles, this analysis is divided according to the subject material of the puzzles. Thus they can be broken down into three sets of puzzles:

1. Understanding depositional environments
2. Understanding fossil sequences
3. Non-geological diachronic thinking (as tested by the "detective puzzle")

Understanding Depositional Environments. Puzzle 2 tested the students' understanding of depositional environments (Figure 4). It required the students to interpret an alternating series of marine and terrestrial strata. A correct interpretation emphasized the fact that there was a temporal change in the level of the sea. The GS were significantly better at making this interpretation and this accounts for much of the significant difference between the two disciplines ($F = 66.58$, df $= 103$, $p < 0.0001$). It is suggested that this difference is due to the GS' participation in fieldwork allowing them to translate environmental indicators (rock and fossil type) into a better understanding of transformational sequences. In addition, an age-related difference was found for this question ($F = 33.94$, df $= 103$, $p < 0.0001$).

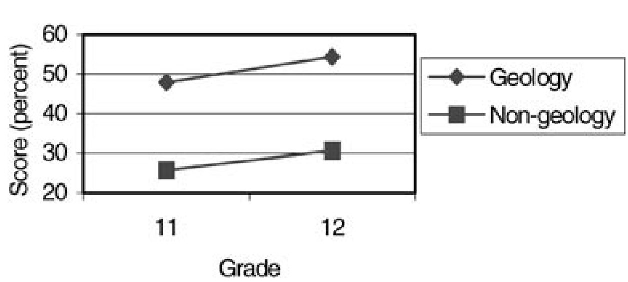


Figure 3. Scores for the GS and NGS on GeoTAT puzzle 1b (depositional environments).

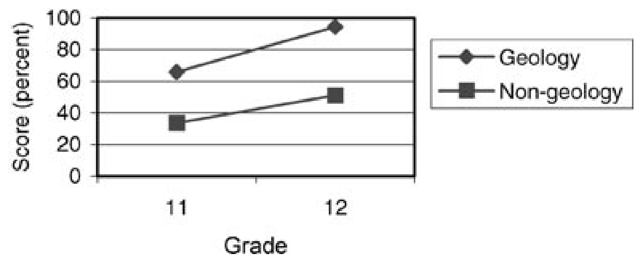


Figure 4. Scores for the GS and NGS on GeoTAT puzzle 2 (depositional environments).

Understanding Fossil Sequences. Three puzzles tested the students' understanding of fossil succession: Puzzle 3a required the students to interpret a sequence of fossiliferous strata as an evolutionary series. There were significant differences between grade 11 and grade 12 students ($F = 5.63$, df = 120, $p < 0.05$); in contrast there was no discipline effect, although the GS outperformed their NGS counterparts (Figure 5) on this puzzle. It might be added that the subject of this puzzle, evolution, is not covered in the earth sciences curriculum; thus, to a degree it does favor students with a background in biology (in this case giving an advantage to the NGS students).

Puzzle 3b required the subjects to determine *why* this same fossilferous series terminated in the upper part of the exposure (Figure 6). Although intended to be an isolated example of causal reasoning, most students treated this question as a diachronic problem, in other words as an extension of the previous puzzle. There were no significant grade or discipline effects in this puzzle. Somewhat surprisingly, the NGS 11 students did better than their GS 11 counterparts, although the GS 12 students were the best performers on this puzzle. It might be added that knowledge of biology (in this case extinction) again might have favored the grade 11 NGS students.

As with the previous puzzle, the subject tested in this puzzle is usually not covered in the geology curriculum. So in essence it might be considered a transfer level problem. Indeed in a test class of twenty-two grade 12 GS who did take part in the implementation of "From Dinosaurs to Darwin" (thus encountering the subjects of evolution and extinction) there was marked improvement from 68.8% (preprogram) to 81.3% after completing the program.

The final fossil sequence (Figure 7) puzzle (Puzzle 6c) in fact fuses two natural cycles together. One cycle consists of a depositional sequence in which an exposure is slowly built of alternating layers of marine and terrestrial rock. The other cycle is based on a series of dinosaur fossils and the remains of their nests, which are found within the terrestrial layers. Taken together, it suggests that the dinosaurs nested in this area during dry periods and

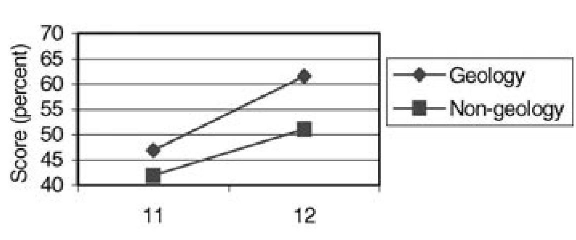


Figure 5. Scores for the GS and NGS on GeoTAT puzzle 3a (fossil sequences).

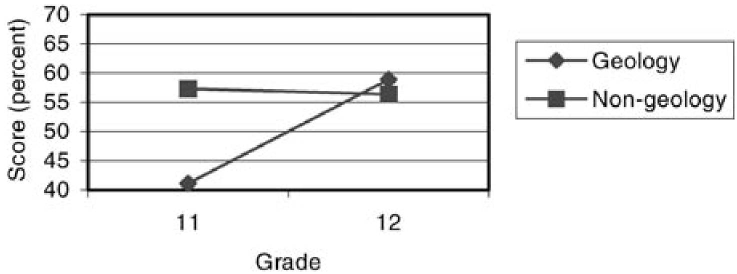


Figure 6. Scores for the GS and NGS on GeoTAT puzzle 3b (fossil sequences).

migrated away during the flood season. (In fact, this puzzle is based on Horner's (1984) description of dinosaur nesting sites in northern Montana.)

No significant differences were found between grade or disciplines. However, a breakdown of this question into biological and geological elements indicates that the NGS were more successful on the former elements and the GS were overall more successful on the latter elements. These results emphasize the often-subtle influence of disciplinary knowledge on diachronic thinking.

Detective Puzzle. Results (Figure 8) from the detective puzzle indicate a significant age difference in the ability to solve the detective puzzle ($F = 7.60$, df $= 137$, $p < 0.01$). It should be remembered that the detective puzzle was designed so as to eliminate domain-specific knowledge. Thus, this age effect seems to contradict Montagnero (1992, 1996) who suggests that by age 11 the full suite of diachronic schemes should be present. This result suggests that the development of diachronic thinking develops into adolescence. In imposing a Piagettian age limit, Montagnero (1996) does not consider strongly enough the child's daily experience with common natural transformations.

The results also indicate that there is no overall discipline effect. However, it is interesting to note that the GS did perform better than their NGS counterparts. This suggests a subtle effect of transfer in which the GS are transferring their diachronic understanding to domains outside of geology. This question however still needs further study.

It might be added that amongst five classes who studied the curriculum, "From Dinosaurs to Darwin" (an integrated unit in paleontology and evolution, which stressed diachronic thinking) there was improvement postprogram in all classes (albeit not significant) on the detective puzzle. This lends support to the argument that exposure to the diachronic approach in geology may very well transfer to domains outside of geology. This has practical

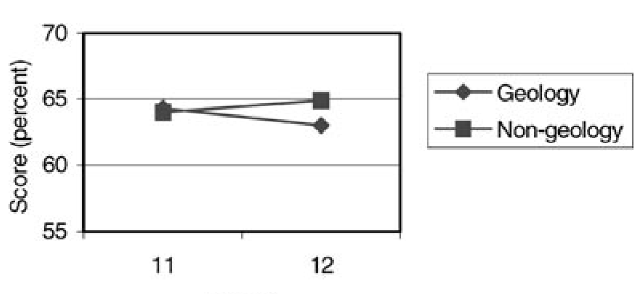


Figure 7. Scores for the GS and NGS on GeoTAT puzzle 6c (fossil sequences).

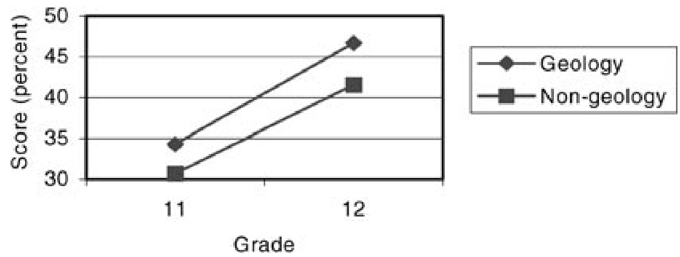


Figure 8. Scores for the GS and NGS on GeoTAT puzzle 7 (detective).

implications for those sciences with an historical perspective, such as ecology, astronomy, or archeology. Although specific knowledge is important for acting upon the diachronic schemes, it is equally important that the students understand the logical schemes upon which diachronic thinking is built.

Puzzles Which Tested the Isolated Scheme of Transformation

This is the only group of questions in which the NGS performed better than the GS, albeit only in the second of these two puzzles (6b) (Figure 10) was there a significant discipline difference ($F = 8.16$, df $= 126$, $p < 0.01$). It should be noted however that these two questions (6a and 6b) (Figure 9 and Figure 10) contained many inferences to biological knowledge that the geology students did not possess. The non-geology students contained a larger population of students who indeed were studying biology and this may have influenced the results.

This suggestion is corroborated by the implementation cycle of the unit "From Dinosaurs to Darwin" (Dodick & Orion, 2000). In this program, students invested considerable time in the biological as well as geological elements of the fossil record. In puzzle 6a preprogram, for example, the grade 12 geology students received 88.9% and improved considerably postprogram to 100%, whereas in puzzle 6b preprogram they only received 37.5% but post-program again improved considerably to 63.0%. In fact, such improvement was the rule amongst all of the students who participated in the implementation of this learning unit. It might be added that many of these students came from non-science backgrounds.

The findings of this study are as follows:

1. The GeoTAT instrument was found to be both a reliable and valid instrument for testing student understanding of geological transformation.

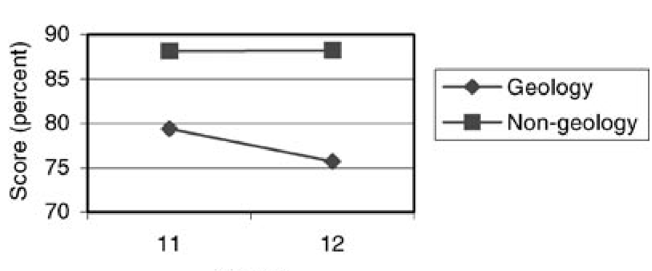


Figure 9. Scores for the GS and NGS on GeoTAT puzzle 6a (environmental interpretation).

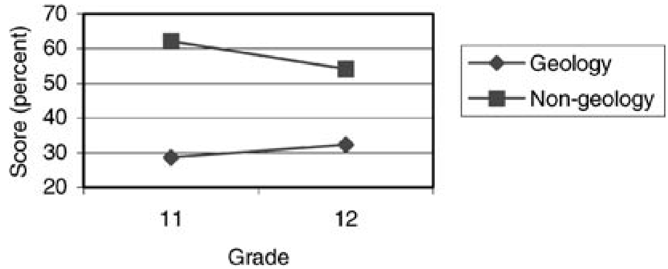


Figure 10. Scores for the GS and NGS on GeoTAT puzzle 6b (sudden extinction).

2. The GeoTAT was sensitive to differences in age as well as discipline (geology or non-geology background). In terms of age, the grade 12 subjects outperformed their grade 11 counterparts on a majority of the GeoTAT puzzles (including most importantly the detective puzzle which is not dependent on discipline-based knowledge encountered in school). This also suggests that the ability to think diachronically continues to improve beyond Montagnero's limit of 11 years in age (Montagnero, 1992, 1996).
3. In terms of discipline differences, when comparing the GS and the NGS, it was seen that the former group held a significant advantage over the later in solving problems in temporal organization and diachronic thinking. This relationship was especially strengthened by the second year of geological study (grade 12) in which the geology majors had accumulated many hours in the field. Field study provided two advantages:

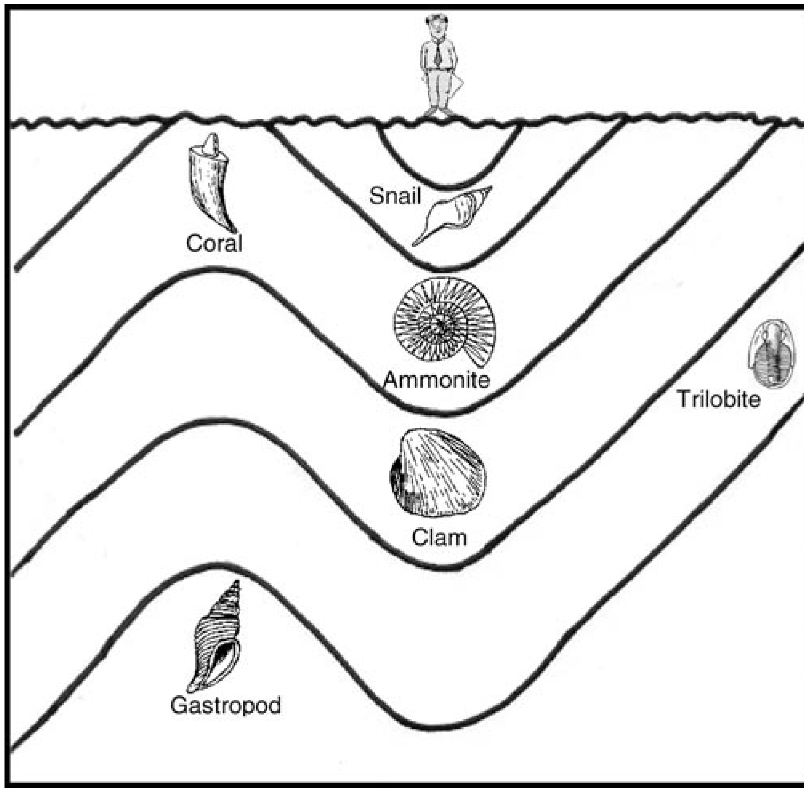
   a. An experience with the 3-D factors influencing an understanding of temporal organization. Rather than experiencing the principle of superposition (for example) just as a simulation or as a 2-D form on paper, the students received experience with real geological structures in the natural environment of their deposition.
   b. Learning about the types of evidence that an earth scientist specifically seeks in understanding a transformational sequence, while concurrently ignoring extraneous evidence that might undermine such an understanding.

4. In terms of difficulty, the easiest of the three sets of geological puzzles to solve for both NGS and GS were those requiring transformational thinking. The most difficult puzzles to solve were the full diachronic puzzles, for the most part because they required the students to use the entire suite of diachronic schemes.
5. It was seen that the geology students were better able to transfer their diachronic thinking skills to other tasks outside of geology. Moreover, with proper training, this skill can be improved. Diachronic thinking is a central skill in any historically based science, such as ecology, astronomy, and evolutionary biology. Thus, students who want to pursue such disciplines should understand their logical basis, which is based on diachronic thinking.

APPENDIX: COMPLETE COPY OF GeoTAT QUESTIONNAIRE

The appendix contains a complete copy of the GeoTAT questionnaire.

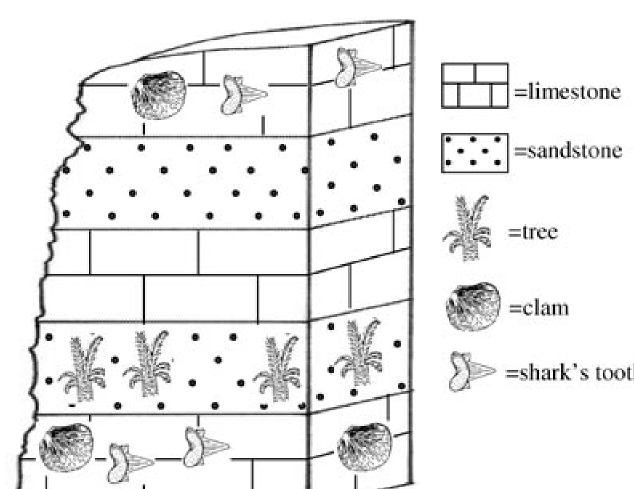
1) The geologist in the diagram below is standing on a column of marine sedimentary rock containing fossils



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).

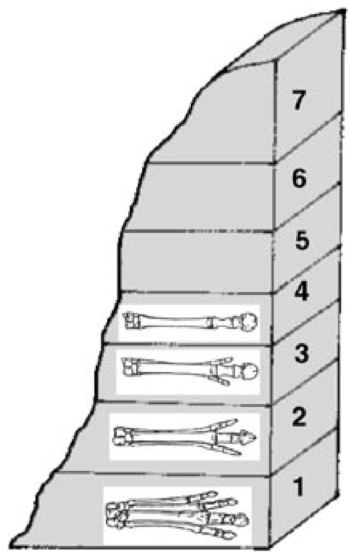
1b) Try to reconstruct the processes in order which lead to the creation of the rock exposure in the picture above.

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.

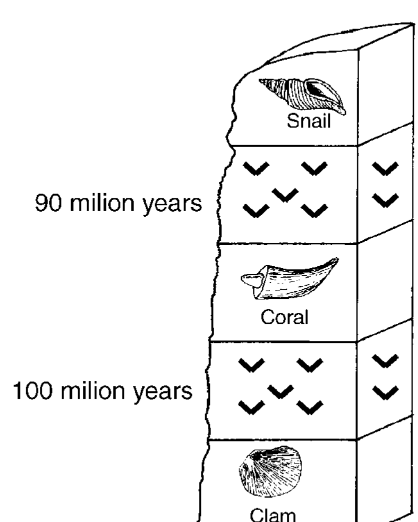
3) The illustration below represents a fossiliferous rock exposure. The fossils are the remains of bones from the feet of unidentified species of mammals.



3a) Try and describe the process that took place between rock layer 1 to layer 4.

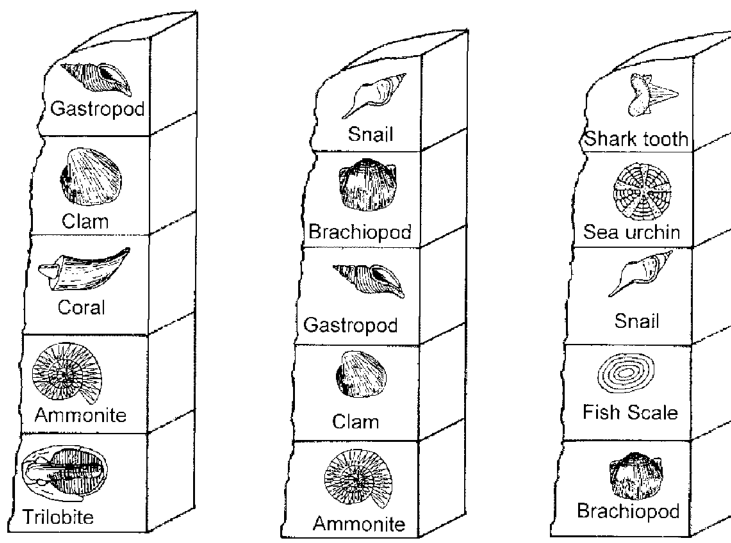
3b) Try to suggest two possible reasons for the absence of fossils after rock layer 4.

4) The following picture represents a rock exposure that contains three types of fossils (snail, coral, and clam). Two layers of igneous rocks (represented by the symbol ✓) lie between the layers containing the fossils. The age of the igneous rock layers have been determined in the lab by scientists.

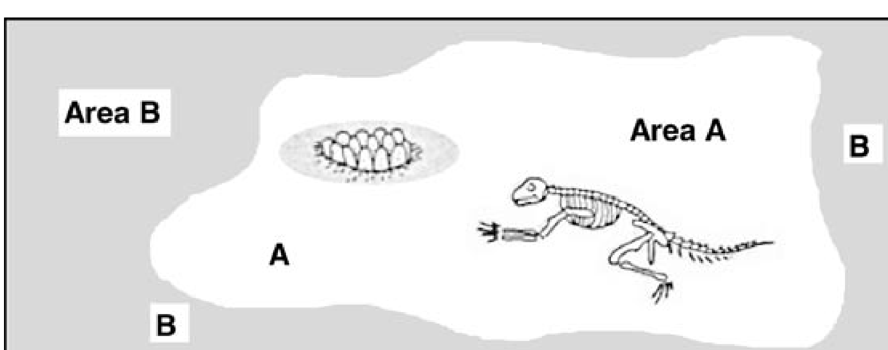


Try to determine the absolute age (in years) of the three different fossils (snail, coral, and clam).

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:

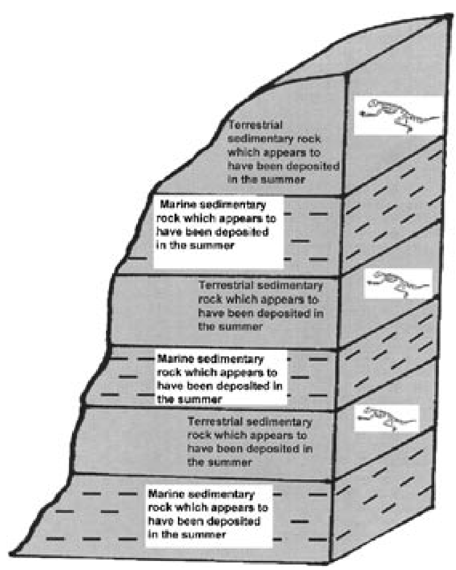
1. The dinosaur skeletons excavated in this area range in size from very tiny to very large. This suggests that in this one area, the age range of the dinosaurs was broad, ranging from newly hatched babies to fully grown adults.
2. In this site, a large series of nests containing fossilized eggs were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

7) Detective Puzzle:
The following is a police report from Chief Inspector Nevo (There is no need to fill in the blanks in this story).

The time was____when the telephone on my desk rang. Over the line I heard a frightened voice. "Rush quickly over to____. It appears as if something terrible has happened in the room of____," he said. "I was just talking to him a minute ago, when suddenly I heard a bang, and then silence descended over the line!"
I immediately drove over to the place to which the caller directed me. I opened the door to the room and peered inside. ____ was lying on the floor. I approached him; he appeared to be unconscious. I nudged his face a bit and noticed a deep cut in his forehead that was covered in blood. I quickly called for medical attention, made sure that no one could enter the room, and ordered the police photographer to take pictures. Meanwhile, my deputees and I began to investigate the circumstances of this incident.

Below is the picture of the crime scene taken by the police. Carefully study the scene in the picture and answer the following questions:

7a) Where exactly, in your opinion, did this incident occur?
7b) Was the “victim” the owner of the room? (Explain your reasoning)?
7c) How and why, in your opinion, was the man lying on the floor injured?
7d) Reconstruct the activities of the “victim” from the moment he entered the room until the time of the incident.
7e) At what time did the incident occur?

REFERENCES

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

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

Bullock, M. (1985). Causal reasoning and developmental change over the preschool years. Human Development, 28, 169–191.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reason. In W. Friedman (Ed.), The developmental psychology of time (pp. 209–205). New York: Academic Press.

Dodick, J. T., & Orion, N. (2000). 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). Geology as an historical science: Its perception within science and the education system. Science and Education.

Everitt, C. L., Good, S. C., & Pankiewicz, P. R. (1996). Conceptualizing the inconceivable by depicting the magnitude of geological time with a yearly planning calendar. Journal of Geoscience Education, 44, 290–293.

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

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

Frodeman, R. L. (1995). Geological reasoning: Geology as an interpretive and historical science. Geological Society of America Bulletin, 107(8), 960–968.

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

Gardner, P. L. (1975). Attitudes to science: A review. Studies in Science Education, 2, 1–41.

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

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

Harbaugh, J. W. (1968). Stratigraphy and geologic time. Dubuque, Iowa: W. C. Brown.

Horner, J. (1984). The nesting behavior of dinosaurs. Scientific American, 250(4), 130–137.

Hume, J. D. (1978). An understanding of geological time. Journal of Geological Education, 26, 141–143.

Koballa, T. (1984). Designing a Likert type scale to assess attitudes towards energy conservation: A nine step process. Journal of Research in Science Teaching, 20, 709–723.

Lyell, C. (1830–1833). Principles of geology (1st ed.). London: Murray.

Marques, L. F. (1988). Alternative frameworks for urban Portuguese pupils age 10/11 and 14/15 with respect to earth life and volcanoes. Unpublished Masters dissertation, Keele University, U.K.

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.

Metzger, E. P. (1992). The strategy column for pre-college science teachers: Lessons on time. Journal of Geological Education, 40, 261–264.

Montagnero, J. (1992). The development of the diachronic perspective in children. In F. Macar, V. Panthas, & W. J. Friedman (Eds.), Time action and cognition (pp. 55–65). Amsterdam: Kluwer.

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

Noonan-Pulling, L. C., & Good, R. G. (1999). 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.

Oldroyd, D. R. (1996). Thinking about the earth: A history of ideas in geology. Cambridge, MA: Harvard University Press.

Orion, N., & Hofstein, A. (1991). The measurement of students' attitudes towards scientific field trips. Science Education, 75(5), 513–523.

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

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

Press, F., & Siever, R. (1998). Understanding the earth (2nd ed.). New York: W. H. Freeman.

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). Misconceptions in the earth sciences: A cross-age study. Paper presented at the 62nd Annual Meeting of the National Association for Researc in Science Teaching, San Francisco, CA.

Spencer-Cervato, C., & Daly, J. F. (2000). Geological time: An interactive team-oriented introductory geology laboratory. Teaching Earth Sciences, 25(1), 19–22.

Trend, R. (1997). 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.

Trend, R. (1998). An investigation into understanding of geological time among 10- and 11-year old children. International Journal of Science Education, 20(8), 973–988.
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(5), 539–555.
Whewell, W. (1837). History of the Inductive Sciences. London: John W. Parker.
Zwart, P. J. (1976). About time: A philisophical inquiry into the origin and nature of time. Amsterdam: North-Holland Publishing.