UNDERSTANDING GEOLOGICAL TIME: A PROPOSED ASSESSMENT MECHANISM FOR BEGINNER AND ADVANCED GEOLOGY STUDENTS AT THE UNIVERSITY OF BRITISH COLUMBIA (VANCOUVER)

by

JAMIL AHMED NIZAM RHAJIAK

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF SCIENCE (HONOURS)

in

THE FACULTY OF SCIENCE

(Geological Sciences)

This thesis conforms to the required standard

……………………………………….
Supervisor

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

MARCH 2009

ABSTRACT

Having a firm grasp of geological time is essential to developing a full understanding of the Earth. Many studies have focused on students in the K-12 and entry-level college education systems. The proposed 20 question, mainly multiple-choice, assessment mechanism is designed to probe the understanding of geological time amongst beginner (entry-level college) and advanced (graduating) students in a major's geology program. A four step process involving: establishing instructor expectations of students, development of an assessment mechanism from existing resources, think-aloud validation with student volunteers, and an iterative refinement process for the developed assessment mechanism revealed insights on student behaviour and creating multiple-choice tests. Student behaviour is assessed via displayed reasoning acts of recalling facts, posing questions, making evaluations, and pausing. From validation interviews students displayed gaps in their understanding of geoscience terminology and a lack of technical vocabulary when reasoning questions out-loud. The refinement process has revealed the following problems associated with developing multiple-choice questions: unclear wording, emphasis of key words, easily eliminated distractors, limitations on cognitive levels of assessment, use of pre-validated questions outside of their context, and testing multiple concepts in one question. The implementation of this assessment should aid in development of the geology curriculum within the Department of Earth and Ocean Sciences at UBC by giving instructors a snapshot of student understanding of geological time. This study serves as a springboard for further scholarly investigations of geology education at UBC.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICIES

LIST OF DIGITAL MEDIA

ACKNOWLEDGEMENTS

From insights on education to scientific writing, this study has been one learning experience after another. I would like to take this opportunity to extend my sincere gratitude to the individuals who helped make this work possible. To Dr. Timothy Heaton (University of South Dakota) for permission to use his existing resource of geological time questions; to Dr. Julie Libarkin (Michigan State University) for early access to the Geoscience Concept Inventory 2.0; to Dr. Wendy Adams (University of Colorado Boulder) for sharing best practices for student validation interviews; and to Dr. Cathy Manduca (Carleton College) for sharing resources and best practices for reporting on geoscience education research.

My appreciation also goes to the Carl Wieman Science Education Initiative for funding my validation interviews and assisting with my ethics certification.

The Department of Earth and Ocean Sciences has been an extremely supportive group of faculty, staff and students. Thank you to the students who volunteered to be involved in the validation process, your willingness to help further teaching in the department is remarkable. Specifically, I would like to acknowledge Dr. Mary Lou Bevier, Dr. Stuart Sutherland, Dr. James Mortensen, Dr. Kelly Russell, Dr. Gregory Dipple, and Dr. Paul Smith for their thoughts on geological time and recognition for the importance of geoscience education research.

Finally, for their unwavering commitment, support, excitement, and patience I would like to thank my project supervisors, Dr. Sara Harris and Francis Jones. Without the weekly meetings for 8 months, nearly instantaneous feedback on submissions, their true passion for improving geoscience education, and so much more none of this study would have been possible. Thank you.

INTRODUCTION

Geological time is a fundamental concept in building a basic understanding of the Earth. Not only is geological time an observation based and testable concept, it has implications for the rates of natural processes and human evolution (Zen, 2001). The understanding of geological time is crucial to a student's mastery of the geosciences (Dodick and Orion, 2003). This study puts forward an assessment to determine if students' abilities to deal with geological time, at different stages of an undergraduate program, match instructor expectations at the University of British Columbia (UBC).

Previous work related to the understanding of geological time has traditionally centered on students in the K-12 system (Trend, 1998; Dodick and Orion, 2003b) and lower-level College (Libarkin et al., 2005; Libarkin and Anderson, 2005). Through these studies there have been a handful of tools developed to assess the abilities of students, including: the Geological Time Aptitude Test, Temporal Spatial Test, Strategic Factors Test (Dodick and Orion, 2003b), the chronoscalimeter (Nieto-Obregon, 2001), geological time conceptest questions (McConnell et al., 2006), and relevant portions of the Geoscience Concept Inventory (Libarkin and Anderson, 2005).

My work specifically aims at addressing an understanding of geological time amongst geology majors at both the introductory level, as reflected by previous work, and at the senior level of study. This sets it apart from other efforts and provides an opening for new directions of research within the UBC context.

With an internal department review of the UBC geology curriculum underway since 2008, now is a good time to evaluate student capabilities. This study begins to evaluate whether experiences in the program provide students with the needed skills to work with concepts that relate to geological time. At a course level, success is based on the interaction between the instructor and the students. The understandings of concepts are gained through lectures, assignments, activities, and examinations (Harris, 2001). At the program level, connections among courses and opportunities to revisit fundamental concepts help determine a student's mastery of geological time when they leave the program.

This study focuses on how well geology students are learning essential concepts related to an understanding of geological time. Foundational concepts identified by instructors in the department include absolute dating, relative dating, Earth history, uniformitarianism, rates and processes, and the geological timescale. Instructors' expectations of students are represented in a multiple-choice assessment tool. Methods followed to create this assessment are outlined in a four step process: establishing instructor expectations, assessment tool development, assessment tool validation, and refinement of questions. The results from this effort are presented as tables of observed student behaviour, areas of improvement for multiple-choice questions, the completed assessment, and the changes made to the assessment throughout the validation and refinement process. Specific examples are used to highlight the results and act as a platform for discussion on my work.

METHODS

The development of this study can be summarized into four steps; establishing instructor expectations, initial assessment development, validation, and refinement (Figure 1).

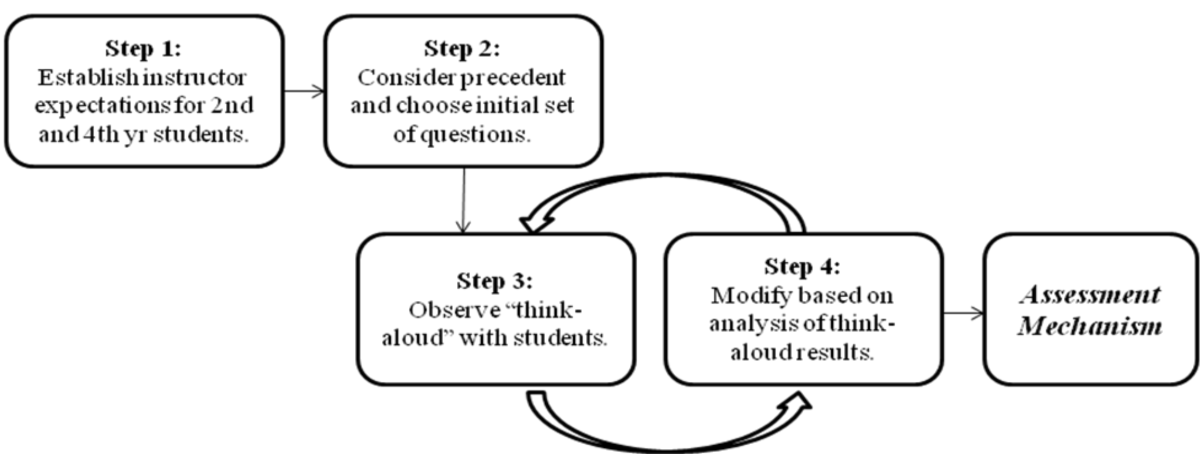


Figure 1. Summary of methods used in this study.

There were many interesting observations and insights made throughout the process that are addressed further in the discussion. The relative simplicity of my methodology is a strength of this work, allowing it to be adapted to contexts outside of UBC.

Step 1: Establishing Instructor Expectations

To obtain an accurate assessment of instructors' expectations of students, I conducted one-on-one interviews with six instructors of core geology program courses. Their expertise includes: petrology, paleontology, sedimentology, fluid-rock interactions, volcanology, tectonics, and isotope geology. Courses within the department that are, or have been, taught by these instructors include: first-year survey courses (general interest and introduction to geosciences), Introductory Petrology, Geological Time, Field Techniques, Sedimentology, Geochemical Thermodynamics, Introduction to Mineral Deposits, Tectonic Evolution of North America, and other upper-level courses. Although students have some flexibility in the geology courses that they take, they will encounter a large number of these instructors throughout their undergraduate careers.

The interviews themselves took approximately thirty minutes of the instructor's time. Each interview was recorded in an audio format with accompanying notes to ensure accuracy of any comments made by the individuals. Questions that were asked of the instructors were designed to extract testable concepts (Table 1).

Table 1. Questions asked of instructors with reasoning.

Question Asked	**Reasoning**
1) How important is the understanding of geological time and processes? Why?	To attain the individuals initial thoughts when geological time is brought up.
2) What specific concepts do you associate with building an understanding of Deep Time?	To attain the concepts that they believe lead to a complete understanding of geological time.
3) When it comes to the previously listed concepts; what do you expect students to do after the first year of their program? (2^{nd} year in Faculty of Science) (ie. What do you test?)	Gauge expectations of students after a year of Geology focused course material (2^{nd} year at UBC).
4) When it comes to the previously listed concepts; what do you expect students to do right before they are finished the program? (about to work or do research)	To explore the instructors idea of what differentiates a graduating student of the program from a student who is just beginning the program with respect to geological time.
5) Any other comments?	To allow for any other thoughts or questions to be addressed.

After the interviews were conducted the next step was to summarize all comments from the instructors. The sub-topics related to geological time that were consistent across all instructors were compared for the different expertise levels. In this study two expertise levels were considered, beginner and advanced. For the purposes of this study, beginners are students who have completed their second year of study in the geology program and, advanced students, are about to graduate from the program. Common expectations of students were combined to create a list of testable concepts (Table 2). Outliers from this process, which were all expectations that were too high of students, were excluded from the testable concepts. Only one instructor had expectations that were significantly higher than those established by his/her colleagues. An example of a high expectation would be for graduating students to know major fossil groups and correlate portions of time that they represent. No outliers were identified at the low-end of instructor expectations.

Table 2. Key concepts and expectations of beginner and advanced students. Italicized font indicates concepts that were not specifically addressed in the final assessment mechanism.

Key Concepts	After 2nd year in program (Beginner)	Before Graduation (Advanced)
Timescale	▪ Understand the immensity of geological time ▪ Familiarity with terminology of geological eons, eras, and periods	▪ Reproduce time scale to eras, periods, and epochs with the associated dates ▪ Explain framework for its construction – that it was based upon succession of fossil types
Relative Dating	▪ Apply knowledge of stratigraphic principles and sedimentary features to create a geological history ▪ Know that fossils can be used to define units because they are separated in time (life changes through time, therefore remains of life can be used to distinguish different periods of time)	▪ Apply relative dating principles to field and map interpretations to create geological histories
Absolute Dating	▪ Recognize commonly used radiometric dating methods ▪ Define the principles of radioactive decay	▪ Explain basic principles of multiple dating processes ▪ Apply multiple dating processes appropriately in different geological settings ▪ *Calculation of ages from data*
Earth History	▪ Describe the history of Earth's formation ▪ Know the relative timing of major geologic events ▪ *Describe plate tectonic theory*	▪ Describe the paleo-geographical development of Earth (influence of tectonic plate movement) ▪ *Reconstruction of Earth history from evidence found in the rock record* ▪ Placement of major geological events on the timescale (extinctions, formations, plate movements etc…) with associated dates
Uniformitarianism	▪ Know the concept and understand the context (technological limitations) in which it was created	▪ *Point to examples of the antiquated nature of the concept*
Rates and processes	▪ *Knowledge of timescales of basic geologic processes (mountain building, volcanism, lava cooling, metamorphic events)*	▪ Quantify geological processes from chemical and physical rate laws

The foundation of this study is instructor input on concepts that they feel are closely tied to geological time. Although not all relevant instructors within the geology program participated in this initial survey there was a representative spread of sub-discipline expertise in geology with the instructors that were interviewed. Each instructor's bias towards their own field of study became apparent once the data was collected. For example, instructors with a background in sedimentology placed emphasis on understanding stratigraphic sequences, fossils as indicators of evolution, and paleo-geographic reconstructions where as instructors with a background in geochemistry placed emphasis on principles of radiometric dating, applicability of different dating methods, and rates of geochemical processes. This example does not outline that each instructor's comments were exclusive to their specific field of study, it only provides observed trends.

<u>Step 2: Initial Assessment Mechanism Development</u>

The assessment mechanism was originally compiled using questions from different sources, including some originals of my own. A main resource was the browsable online database of *Example ConcepTest Questions* (McConnell et al., 2006), a part of the *On the Cutting Edge - Professional Development for Geoscience Faculty* module from the Science Education Resource Centre at Carleton College (SERC). Other sources included the Geoscience Concept Inventory v.2.0 (Libarkin, personal communication, 2008), and a geological time practice exam written by Dr. Timothy Heaton, Professor of Earth Science at the University of South Dakota.

Much thought was put into the method that would be appropriate to use in the assessment mechanism. Two criteria that had to be met were (1) easy implementation and (2) simple scoring to produce results that could be interpreted in a consistent manner for curriculum planning and program assessment. Multiple-choice tests fit these criteria because they allow many concepts to be addressed in a short period of time and allow for rapid scoring of results (Burton et al., 1991). For this assessment the anticipated completion time of the test is 30 minutes.

Not all of the key concepts expressed by instructors (Table 2) were covered by the set of questions that I used. Reasons for this exclusion stem from not being able to formulate questions that would provide a sufficient degree of difficulty for students or inadvertently testing multiple concepts when trying to increase the degree of difficulty for a question. These points are addressed in the discussion of this work.

The final product is one test with a combination of twenty multiple-choice and written answer questions ranging in difficulty. Questions testing lower cognitive levels of understanding like knowledge and comprehension are mixed with questions testing higher cognitive levels like application, analysis, and synthesis (Bloom, 1956; Appendix 1). The cumulative result of these questions and the following validation process is a mechanism that assesses students on many of the key concepts outlined by UBC instructors (Appendix 2).

Step 3: Assessment Mechanism Validation

The purpose of validation was to ensure that questions in the assessment were understood by students in the intended manner. A common problem with un-validated assessments is that questions may not test the skills an instructor intended simply because many students misinterpret the question. In these cases an instructor does not learn his or her students' true capabilities. Because the proposed set of questions has been validated, any results from the test can accurately be interpreted (Sireci, 2007) and used for enhancing the UBC geology curriculum.

Student volunteers were solicited by department-wide emails and posters. Volunteers had to meet simple criteria; they must have been students in the geology or the earth and ocean sciences majors programs and have taken the EOSC 222: Geological Time course or an equivalent course at another institution. As per Carl Wieman Science Education Initiative (CWSEI) guidelines on student interviews each student signed a standard consent form (Appendix 3), received a $15.00 compensation for their time, and the guarantee that any personally identifiable information would be kept in confidence.

The validation process involved 30 to 45 minute private sessions with ten volunteer students and the principle investigator. Table 3 provides a breakdown of gender and year of study of the participants.

Table 3. Student participant demographics: gender and year of study.

	Gender		**Year of study**	
	Male	Female	Year 3	Year 4 or greater
n	4	6	2	8

The experience of other investigators indicates that little new information is yielded after 8 – 10 interviews (Adams, personal communication, 2009) the small number of student volunteers, 10, can still be taken as acceptable. This is echoed in the development of question 3 (Appendix 1) where students are asked to create a geological history from a cross-section presented in question 2. Each student clued in on the discrepancy of the erosion feature in the cross-section through multiple versions of the assessment. No new information relating to the clarity of the figure was gained other than to define the erosion feature, which has been done in question 2 of the final assessment (Appendix 2).

Students were instructed to work through the multiple-choice testing mechanism while verbalizing all of their thought processes, similar to Norris's "Think Aloud" elicitation level (1990). While students were engaged in the test they were not interrupted by the investigator or their surroundings. This included the denial of requests for the investigator to answer questions related to the test. In addition, voice recordings of the sessions were made to ensure accuracy of the students' interpretations of questions. Due to confidentiality agreements (Appendix 3) copies of the audio recordings are available upon request of the author, however, digital copies of tests taken by students are available (Digital Media 1).

Think aloud validation was used for two principle reasons: first, think aloud validation without investigator interruption is seen as a way to supplement the inherent weakness of multiple-choice tests to tease apart critical thinking processes (Norris, 1990), and secondly, it "parallels the 'free report[ing method]' that yields the most accurate eyewitness testimony" (Loftus, 1979).

Step 4: Assessment Mechanism Refinement

After every session, each question of the assessment mechanism was analyzed for potential weaknesses that are discussed later in this paper. Questions were added, altered, or deleted to address areas that required improvement. Iterations of the assessment mechanism received unique version numbers, creating a simple method to track changes (Appendix 1). The final version of the assessment tool has been altered from student validations of version 5 and has been formatted for consistency (Appendix 2).

An example of the validation process reveals problems with student interpretation of questions and the steps taken to improve its clarity (Figure 2.). The draft of the question, which is from the Geoscience Concept Inventory, was passed by two department instructors who provided initial feedback. This input is reflected in Version 1 of the assessment. Student validation revealed that options (A) through (E) did not accurately reflect the age of the Earth, 4.5 billion years, but that option '(d)' was approximately the correct answer. The question was altered to ask students to select the option which approximately gave the correct answer. Version 2 of the question (Figure 2) was carried through from Versions 2 through 6 of the assessment.

One question of a validation session was deemed inadmissible as the student being interviewed revealed that he/she had a mild learning disability. The students' interpretation of the omitted question was clearly impacted by the disability. The rest of their validation interview was not affected in any apparent way and is included in the participant count (Table 3). This test has not been created in consultation with UBC's Access and Diversity office and does not exist in alternate formats for individuals with disabilities.

Figure 2. Validation of question 1 from the proposed assessment mechanism.

Draft	**Instructor Feedback**	**Version 1**	**Information from validation**	**Version 2**
If you could travel back in time to when the Earth first formed as a planet: How many years back in time would you have to travel? (a) 4 hundred years (b) 4 hundred-thousand years (c) 4 million years (d) 4 billion years (e) 4 trillion years	▪ Is a preface and proposing time travel necessary for students majoring in geology? – Most likely not.	1) How many years back in time did the Earth form? (a) 4 hundred years (b) 4 hundred-thousand years (c) 4 million years (d) 4 billion years (e) 4 trillion years	▪ Student indicated that answers do not reflect the specific age of the earth, but that the correct answer is 'D' because it is approximately correct.	1) Approximately how many years back in time did the Earth form? (a) 4 hundred years (b) 4 hundred-thousand years (c) 4 million years (d) 4 billion years (e) 4 trillion years

RESULTS

The product of this study is a twenty-question assessment that addresses most of the key concepts identified by faculty members (Appendix 2). After six iterations, achieved through the validation process, the result is nineteen multiple-choice questions and one written question.

The validation process revealed a number of expected reasoning acts (Norris, 1990) plus unexpected observations of student behaviour (Table 4). Validation also made several areas of weakness apparent in the questions that were being asked (Table 5).

The final version of the assessment has had each question correctly interpreted by students. While students did not correctly answer all the questions during validation it was clear that they knew what was being asked.

Table 4. Expected (Norris, 1990) and unexpected observations in student behaviour.

Expected observations from Norris (1990)	
Reasoning Act	**Behaviour**
1) Citing factual details	Either recalling a factual detail given in an item prior to the one currently being done, recalling such a prior detail incorrectly, or stating a detail in the current item
2) Self-questioning	Posing questions that appear to be directed to the subject rather than to the interviewer
3) Making evaluations	Either evaluating previously stated judgments or conclusions, or evaluating ones that had not been verbalized
4) Pausing	Either making verbal inflections (Ohhh! Mmmm!), or being silent.
Unexpected observations from this study	
Observation	**Description**
5) Lack of descriptive vocabulary	Student thought process uses lay terms to describe units, relationships, or specific items (“Thingy”, “Between these other things”)
6) Terminology gaps	Lack of knowledge of simple geosciences terms (geological history, index fossil, accretion)

Table 5. Descriptions of common weaknesses in questions.

Common weakness	**Description**
1) Unclear wording	Questions were worded awkwardly or did not focus on a specific outcome required by the student
2) Testing multiple concepts	Questions were testing the understanding of more than one concept (Knowledge of order and specific dates on the time scale)
3) Key-word emphasis	Words that were important in the correct interpretation of a question were not highlighted (“best”, “not”, “primarily”)
4) Inappropriate distractors	Distractors that could be eliminated too easily or were not clear when interpreted by students
5) False positives	Questions that allow for the correct answer to be reached by incorrect reasoning

DISCUSSION

How Students Think

Version 1 of the testing mechanism contained slightly altered questions from the resources mentioned above. An initial revision of the questions with two instructors was carried out to point out any questions that did not match the cumulative expectations of the instructors that were interviewed. Through the validation process several of Norris's 'reasoning acts' (1990) became apparent and are outlined in Table 4. These are useful because they highlight ways that students interpret questions.

Citing factual details

Reasoning act 1, Citing factual details, occurred multiple times per interview. Students would recall various facts such as the order of Periods on the geological timescale, uses of radiometric dating methods, and specific ages of major geologic events. This can be attributed to a majority of the questions that were being asked requiring factual recollection to correctly answer them. When interviewing students it was clear, with questions that tested basic principles, that they either immediately knew the answer to a question or spent some time trying to recall a fact in place of working through the question using basic principles and information that was provided.

Self-questioning

Reasoning act 2, Self-questioning was inferred by students' repetition of the proposed questions or parts of the questions and blank stares at the test. When they did not know an answer to a question they tried to rephrase it to evoke factual recollection. This process is clear when a student was answering question 17 of version 5 of the assessment, "…let me think of what the Triassic was like…I think in the Triassic it was not very cold…" If the rephrasing of a question did not work, students would spend time waiting to see if they could remember the correct information.

Making evaluations

Reasoning act 3, Making evaluations, was apparent with all of the students that were tested. Each student, at several points throughout their interview, evaluated each distractor of particular questions and as a result either eliminated them as incorrect or grouped them as potentially correct answers. This is clear when a student working through question 19 of version 2 of the assessment (Appendix 1) stated, "'I-J' seems too close together so it's wrong, 'A-H'…I guess that's possible…'H-J', I think that's too short…'E-I'…I don't know. I'm gonna go with C." Similar to reasoning act 2, this action is a test taking skill that students find effective when facing knowledge gaps. Students would often select answers that they felt 'sounded' correct as opposed to ones that they knew were correct.

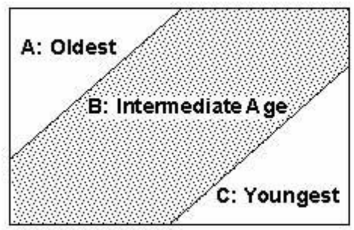
Pausing

Reasoning act 4, Pausing, often followed reasoning act 2 as students processed answers to their self-posed questions in silence. One of the drawbacks to using a think aloud validation method is the inability to ask probing questions because any silent thought remains inaccessible to the research.

The art of writing multiple-choice questions (Vale, 1995) is not easy to master. Throughout the validation process it was continually noted that students were tested on their knowledge of specific topics and not their ability to consume, evaluate, and apply information. I attempted to include questions that would target higher cognitive levels such as question 3; the written response question (Appendix 2), discussed in Strengths and Weaknesses of Multiple-Choice Questions, and question 6 (Figure 3).

Question 6 was intended to determine if a student could make the best inference after evaluating a cross-section and applying their knowledge of stratigraphic principles. Through validation sessions it was clear that many students were not taking their inferences to the furthest and most plausible explanation of the diagram, instead they would stop at '(e).' In this case the best answer is '(a).' Even though '(e)' may be true it does not provide the most useful information from the sketched observation.

6) During fieldwork in western Canada, an experienced geologist sketched the cross section below showing three different units of tilted rocks and their relative ages. What could you **best** infer from this diagram?



(a) Mountain building has overturned the rock units.
(b) Rock layer C formed first and so lies on the bottom.
(c) Rock layers B and C must be igneous rocks.
(d) The sequence of rock layers conforms to the principle of superposition.
(e) The sequence of rock layers conforms to the principle of original horizontality.

Figure 3. Question 6 from the final assessment mechanism.

Terminology gaps & lack of descriptive vocabulary

The re-working of questions to target higher levels of understanding and knowledge was significantly slowed when it was revealed that many upper level students had difficulty understanding basic geoscience terminology such as “index fossil”, “geological history”, and “accretion”. This shows that the use of geoscience terms in exam questions may simply be testing vocabulary when the intended outcome is to test understanding.

Terminology gaps can be coupled with the use of lay terms to describe geological processes. Students often referred to representations of fossils in question 8 (Appendix 2) as “swirly things” or relationships between units in questions 2 and 3(Appendix 2) as “this thing is in between the other two things.” This behavior may be more indicative of a test taking strategy, as opposed to a problem with student understanding.

<u>Cognitive Levels of Assessment Questions</u>

Each question of the final assessment has been analyzed in terms of Bloom’s Taxonomy (1956) to give insight into the level of student understanding being tested (Appendix 1). The spread of the cognitive levels of knowledge, comprehension, application, and analysis allows us to determine the main outcomes of implementing this assessment (Table 6).

Table 6. Summary of cognitive levels of understanding in the final assessment mechanism.

Cognitive Level of Understanding	**Questions testing Level**	**Frequency**
Knowledge	1,5,7,10,12,16,17,18	8
Comprehension	4,13,19	3
Application	2,6,8,14,15	5
Analysis	3,9,11,20	4

From the summary of levels of understanding (Table 6) it is clear that the assessment is dominated by questions testing knowledge, the most straight forward and precise type of question to implement (Bloom, 1956). The implication of this bias towards knowledge based questions is that the assessment tests mainly for students' knowledge of instructor defined abilities of beginner and advanced students.

The two highest cognitive levels as defined by Bloom (1956); synthesis and evaluation, are not represented for any geological time related concepts. These upper levels of understanding may allow for a true separation between the students who have just learned concepts and simple applications of them and the students who have the ability to synthesize new information from their understanding of basic principles and make judgments based on information provided.

Strengths & Limitations of Multiple-Choice

There are clear advantages of multiple-choice testing; however, there were barriers to developing the assessment. One example is the inability to ask students to organize their thoughts or to do a specific task (Burton et al., 1991). For example, it was difficult to construct a multiple-choice question that tested a student's ability to produce a geologic history from a cross section. I have begun to work around this issue by incorporating a single short-answer response, question 3, which asks students to produce a point form sequence of events that would produce a cross section based upon a figure (Appendix 1). This addition requires more time for the scoring procedure of this test but it does provide analysis of a student's ability to use temporal organization, a diachronic scheme that allows students to understand change through time (Montanegro, 1996). In geology this correlates with using principles of superposition, correlation, and original horizontality to establish relationships amongst strata (Dodick and Orion, 2003b). Another benefit of using this question is that it requires the ability to address spatial and temporal understanding, a basic and valuable skill of geologists (Black, 1991).

Another disadvantage to using a multiple-choice testing mechanism is the potential to create false knowledge. Students who take a test can sometimes leave the test with the impression that the presented distractors are true (Roediger and Marsh, 2005). I believe that this effect may be significant for a learning environment where multiple-choice testing is being implemented to supplement instruction. In this study the purpose is to create a snapshot of students' understanding of geological time, we can therefore consider the false knowledge effect negligible. In fact, it is advised that the proposed test not be used as a teaching aid because that will prove ineffective as a method of furthering a students' understanding of geological time.

In addition to the advantages presented in the Methodology section, validating a multiple-choice test allows the development of better distractors. Instructor-generated distractors that are unappealing can be replaced by those that reveal true student misconceptions rather than what instructors believe students may be thinking.

Solutions to Common Weaknesses in Questions

The reasoning acts (Table 4) were indicators of student behaviour. Indicators of weaknesses in questions included direct observation of portions of questions that were problematic and inferred problems based on students' difficulty with answering the questions (Table 5).

Unclear wording

Unclear wording of questions was corrected by simply re-wording the question. Either removing extraneous information or clarifying definitions to make the question as easily understood as possible was carried out based upon information from student validations. An example of the re-wording of a question is clear in comparison of question 7 between version 1 and version 6 of the assessment (Table 7).

Table 7. Comparison of question 7 between version 1 and version 6 of the assessment.

Version 1	**Version 6**
7) Which answer best describes what the surface of the Earth would be when the Earth first formed as a planet? (a) The Earth's surface was covered with jungles (b) The Earth's surface was covered with water (c) The Earth's surface was similar to modern undeveloped areas (d) The Earth's surface was covered with melted rock (e) The Earth's surface was covered with ice	7) Which answer **best** describes what the surface of the Earth was like when the Earth first accreted as a planet? (a) The Earth's surface was covered with solid rock (b) The Earth's surface was covered with water (c) The Earth's surface was covered with solid continental masses (d) The Earth's surface was covered with melted rock (e) The Earth's surface was covered with melted iron

The specific substitution of the word 'formed' with 'accreted' in question 7 (Table 7) came from students being confused with the stage of planetary formation the word 'formed' was referring to. 'Accreted' solved this issue as it refers to the point at which the planet had not yet differentiated and the correct answer '(d)' would be clear, except to those who did not have an understanding of planetary formation.

Key-word emphasis

Students frequently circled or underlined specific words in questions. In later iterations of the mechanism I decided to emphasize these words and phrases using bold and underline fonts. Specific examples include the key-words, "approximately", "best", and "compare". Many questions in the final version of the assessment show this including question 7 (Table 7).

Inappropriate distractors

Other identifiable problems were the use of inappropriate distractors. Some initial versions of questions contained distractors that were easily eliminated without any specific knowledge of geology. Distractors in questions 5 (Table 8) and 7 (Table 7) were changed between versions 1 and 6 of the assessment.

For question 5 it was clear that the distractors '(a)', '(b)', '(d)', and '(e)' from version 1 could easily be eliminated by knowing that radiometric dating became available to geologists after the time scale was established. The distractors were changed to be more historically plausible and difficult to eliminate.

Table 8. Comparison of question 5 between version 1 and version 6 of the assessment.

Version 1	**Version 6**
5) Which method was primarily used to establish the Geologic Time Scale? (a) Carbon-Nitrogen (b) Potassium-Argon (c) Relative dating (d) Rubidium-Strontium (e) Uranium-Lead.	5) Which method was **primarily** used to **establish** the Geologic Time Scale? (a) Correlation of magnetic signatures in rocks (b) Calculation of alpha decay of isotopes (c) Calculation of beta decay of isotopes (d) Correlation of fossils in rock units across vast distances (e) Correlation of rock types across vast distances

Question 7 (Table 7) is another example of inappropriate distractors being used. Options involving jungles, modern undeveloped areas, and ice were also too easily eliminated by students majoring in geology. These distractors were modified to reflect more plausible answers associated with the formation of the Earth. In later validation sessions students had a more difficult time answering this question correctly.

It was surprising to see distractors that came from the Geoscience Concept Inventory, like the ones in question 7 (Table 7), being easily eliminated by students. This shows the importance of the context for which a test is designed. GCI questions are considered thoroughly validated but they are meant to be, and were, implemented in introductory geoscience courses (Libarkin and Anderson, 2005). This differs from the context of this study which is to assess levels of knowledge and understanding in both beginner and advanced students. The need to modify a previously validated question may also be attributed to the high number of upper level, “advanced”, students involved in the validation process; which points back to the context of this study and that it was not designed for beginners unlike other work done in this field (Libarkin et al., 2005). This is a clear indication to instructors that there can be difficulties if previously validated questions are used outside of their original contexts.

Another link to poor distractors could lie in the number of distractors that were used. Each question was designed to have five distractors to be consistent with most of the testing that is done within the department and the various online sources of questions. Efforts to add distractors that were consistent and relevant to questions (Burton et al., 1991) may have compromised the integrity of the questions as a whole. While the optimal distracter count of 3, as some research has pointed to (Haladyna and Downing, 1993), may not be sufficient to test students when a question has many plausible answers, it is clear that developing distractors appropriate for geology majors is not an easy task.

Testing multiple concepts

Testing multiple concepts stemmed from attempting to create questions that tested higher levels of understanding. The results were questions that combined specific knowledge of two concepts instead of an understanding of either concept. Question 12 from version 5 of the assessment (Figure 4) tests whether students are familiar with the time scale and have specific knowledge of the formation of the Himalayan Plateau.

12) Mt. Everest, along with the rest of the Himalayan Plateau began forming…

(a) In the Middle Proterozoic when India drifted Northward into Eurasia.
(b) In the Late Cretaceous when Eurasia drifted Southward into India.
(c) 70 Million years ago when India drifted Northward into Eurasia at a divergent boundary.
(d) 70 Million years ago when India quickly moved Eastward into Eurasia.

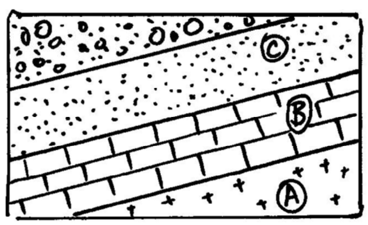
Figure 4. Question 12 from version 5 of the assessment mechanism.

From student validations it was clear that students either knew the movement of the tectonic plates and guessed between options '(a).', '(c).', and '(e)' or knew the time of the event and guessed between options '(b).' through '(e).'. The act of a student choosing an incorrect answer would not identify which portion of the distractor the student did not know. This specific question was removed from the final version of the test.

False positives

False positives during validation were not frequent problems but are noteworthy. They mainly occurred in the initial iterations of the test. The combination of unclear wording and inappropriate distractors allowed for students to arrive at correct answers through incorrect reasoning or knowledge. A false positive occurred during the validation of question 11 in version 1 of the assessment (Figure 5).

On their paper copy of the test the student drafted out a sketch of the geological time scale and began to label it. He/She successfully labeled the Permian-Triassic boundary but incorrectly labeled the Periods of the Mesozoic as Periods of the Paleozoic. The student quickly eliminated the Cambrian distractor as they associated it with being at the beginning of the time scale. He/She proceeded to guess the correct answer as 'Devonian' because they knew that it was below the Permian-Triassic boundary. From a clear gap in knowledge of the time scale the student was able to guess the correct answer. This is due to distractors that were major, recognizable, sections of the time scale. In the final version of the test, question 11 (Appendix 1) has distractors that are temporally closer on the timescale to the correct answer, '(c)'. This is to test specific knowledge of the order of the time scale at a Period level in the Paleozoic.



11) Layer A is from the Ordivician, Layer C is from the Carboniferous. What geological time Period is most likely represented by Layer B?

(a) Paleozoic
(b) Cretaceous
(c) Devonian
(d) Cambrian
(e) Mesozoic

Figure 5. Question 11 from version 1 of the assessment mechanism.

The way students apply their knowledge of geoscience concepts at an advanced level of study may prove to be an important future investigation. It is known that diverse and engaging instruction methods provide a higher level of memory retention (Nuhfer and Mosbrucker, 2007) and despite this study not incorporating an implementation of the assessment it does suggest that assessments targeting instructor expectations can yield useful information on student knowledge and thinking.

Future work

From this study come several areas which deserve continued investigation. This should serve as a solid spring-board for geoscience education research within the Department of Earth and Ocean Sciences at UBC.

Implementation

The implementation of the proposed assessment is a key step in evaluating students' understanding of geological time. It is designed to be implemented with students majoring in geology who are completing their second year in the Faculty of Science as well as the cohort of geology majors who are graduating. Results on specific questions will indicate those students who are capable of higher level processes against students who are limited in their understanding or knowledge of geological time.

Assess remaining instructor expectations

It is clear that developing questions to test all the instructor outlined geological time concepts was difficult. This resulted in several expectations of beginner and advanced students to be left out of the final assessment (Table 2). For example, finding meaningful ways to test plate tectonic theory and the antiquated nature of Uniformitarianism in the context of geological time would make this a much more robust assessment tool.

Continued validation and improvement

Having questions continually validated will offer better insights into students' problem-solving processes and, ultimately, more precise questions (Serci, 2007) that focus on testing understanding versus knowledge. Quantitative analysis of student reasoning acts (Norris, 1990) may also provide a more robust data set on student understanding of geological time than observational evidence can provide alone.

Adding and validating questions to assess higher cognitive levels of thinking would help eliminate the assessments bias towards testing knowledge over other skills such as synthesis or evaluation.

Refinement as curriculum evolves

The ultimate goal of this project is to assist with the development of the UBC geology curriculum. As the curriculum continues to develop, it would be prudent for any future version of this assessment to accurately reflect those changes. This would involve revisiting instructors' expectations as needed and conducting a more inclusive survey of all the instructors within the geology program.

Potential for pre and post testing

The intended implementation for the proposed assessment would provide a snapshot of student understanding within the department at the time of deployment. Consistent deployment of this assessment, or some future version of it, would track how students' understanding of geological time evolves throughout their undergraduate careers.

CONCLUSION

The assessment presented is the result of a four step process involving: instructor input, test creation, think-aloud student validation, and an iterative approach to test refinement. It targets the fundamental concept of geological time and how both beginner and advanced students internalize it.

This study shows that, when tested with the proposed assessment, students demonstrate predictable 'reasoning acts' such as recalling facts, posing questions, making evaluations of test items, and pausing. The significance of these acts is that they give insight into how students process the concepts related to geological time and the questions in which these concepts are tested. Surprisingly, the validation process also revealed that students lack some basic geoscience vocabulary. This is valuable for instructors who are designing assessment questions because it reveals that assumptions about students' understanding of vocabulary may lead to questions that do not produce data on student understanding of geoscience concepts.

Common issues with constructing multiple-choice questions have also arisen. These difficulties include limitations on the cognitive levels of assessment, unclear wording, emphasis of key words or phrases, the use of easily eliminated distractors, testing more than one concept, and false positive responses from students. Testing multiple concepts in one question reveals no conclusive evidence as to what a student does not understand and is another crucial point to be aware of when creating assessment questions.

The use of existing geoscience education resources outside of their intended contexts – which is mainly to test beginner students - may have contributed to some of the previously listed issues. Many of the problems with the pre-existing questions were resolved or at least made apparent via the validation and refinement processes.

This is an initial step in creating a scholarly approach to assessing student understanding of geological time within the Department of Earth and Ocean Sciences at UBC. Subsequent steps will need to address issues of questions that test higher cognitive levels, other than knowledge, as well as the continued development of the test items through validation (Serci, 2007). I look forward to seeing what studies and changes in the department come out of this work.

WORKS CITED

Black, A.A., 2005, Spatial ability and earth science conceptual understanding, Journal of Geoscience Education, v. 53, p. 402-414.

Bloom, B.S., editor, Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners: Handbook I: Cognitive domain, New York, David McKay Company, Inc., 207 p.

Burton, S.J., Sudweeks, R.R., Merrill, P.F., and Wood, B., 1991, How to Prepare Better Tests: Guidelines for University Faculty, Brigham Young University Testing Services and The Department of Instructional Science. http://testing.byu.edu/info/handbooks/betteritems.pdf

(a)Dodick, J., and Orion, N., 2003, Measuring student understanding of geological time, Science Education, v. 87, p. 708-731.

(b)Dodick, J., and Orion, N., 2003, Cognitive factors affecting student understanding of geologic time, Journal of Research in Science Teaching, v. 40, p. 415-442.

Haladyna, T.M., and Downing, S.M., 1993, How many options is enough for a multiple-choice test item?, Educational and Psychological Measurement, v. 53, p. 999-1010.

Harris, M.T., 2001, Strategies for implementing pedagogical change by faculty at a research university, Journal of Geoscience Education, v. 49, p. 50-55.

Libarkin, J.C., Anderson, S.W., Science, J.D., Beilfuss, M., and Boone, W., 2005, Qualitative analysis of college students' ideas about the earth: interviews and open-ended questionnaires, Journal of Geoscience Education, v. 53, p. 17-26.

Libarkin, J.C., and Anderson, S.W., 2005, Assessment of learning in entry-level geoscience courses: results from the geoscience concept inventory, Journal of Geoscience Education, v. 53, p. 394-401.

Loftus, E.F., 1979, Eyewitness testimony, Cambridge, MA, Harvard University Press, In: Norris, S.P., 1990, Effect of eliciting verbal reports of thinking on critical thinking test performance, Journal of Educational Measurement, v. 27, p. 41-58.

McConnell, D.A., Steer, D.N., Owens, K.D., Knott, J.R., Van Horn, S., Borowski, W., Dick, J., Foos, A., Malone, M., McGrew, H., Greer, L., and Heaney, P.J., 2006, Using conceptests to assess and improve student conceptual understanding in introductory geoscience courses, Journal of Geoscience Education, v. 54, p. 61-68.

Montagnero, J., Understanding Changes in Time. London, Taylor and Francis.

Nieto-Obregon, J., 2001, Geologic time scales, maps, and the chronoscalimter, Journal of Geoscience Education, v. 49, p. 25-29.

Nuhfer, E., Mosbrucker, P., 2007, Developing science literacy using interactive engagements for conceptual understanding of change through time, Journal of Geoscience Education, v. 55, p. 36-50.

Roediger, H.L., III, and Marsh, E.J., 2005, The positive and negative consequences of multiple-choice testing, Journal of Experimental Psychology: Learning, Memory, & Cognition, v. 31, p. 1155-1159.

Norris, S.P., 1990, Effect of eliciting verbal reports of thinking on critical thinking test performance, Journal of Educational Measurement, v. 27, p. 41-58.

Sireci, S.G., 2007, Comments on Lissitz and Samuelsen: on validity theory and test validation, Educational Researcher, v. 36, p. 477-481.

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

Vale, D.C., 1995, Book review: developing and validating multiple-choice test items by Thomas M. Haladyna Hillsdale NJ: Erlbaum, 1994, 227 pp., $49.95, Applied Psychological Measurement, v. 19, p. 205-207.

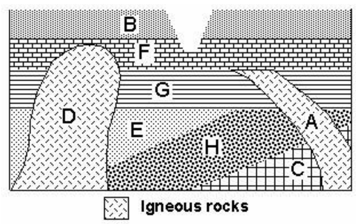
Zen, E., 2001, What is deep time and why should anyone care?, Journal of Geoscience Education, v. 49, p. 5-9.

[This page has intentionally been left blank.]

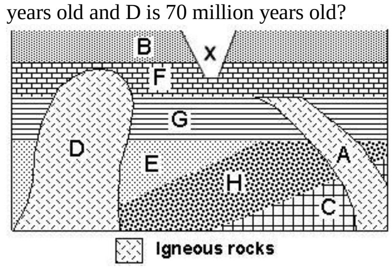
Appendix 1. Changes made between the draft and final versions of assessment questions, instructor identified concepts, and Bloom's cognition level of questions.

[Note: This appendix has been formatted to span across a printed and bound copy of this thesis]

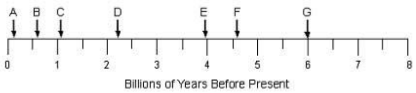
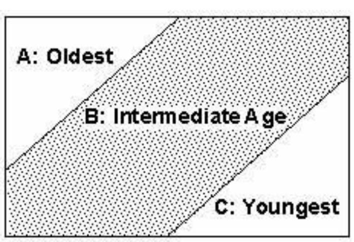
Appendix 1. Continued…

Draft of Question & Source	Key concept from Table 2.	Bloom's Cognition Level	Changes in v1	Changes in v2	Changes in v3
1) **If you could travel back in time to when the Earth first formed as a planet:** How many years back in time would you have to travel? (a) 4 hundred years (b) 4 hundred-thousand years (c) 4 million years (d) 4 billion years (e) 4 trillion years Source: Geoscience Concept Inventory	**Earth History** – Describe the history of Earth's formation	Knowledge	-Remove preface and notions of time travel – unnecessary for geology majors -Add the word "approximately" in the question after students indicated no answers were exactly correct		
2) What is the best estimate of the age of F if A is 100 million years old and D is 70 million years old?  (a) 85 Myrs (b) 170 Myrs (c) 55 Myrs (d) 110 Myrs (e) not enough information provided (f) same age as A (g) same age as D Source: http://serc.carleton.edu/2975	**Relative Dating** – Apply knowledge of stratigraphic principles and sedimentary features to create a geological history	Application		-Re-ordered options to be in numerical order (Burton et al, 1991) -Removed "not enough information provided" distractor (Burton et al, 1991)	-Addition of erosion feature label "X" on diagram in response to validation of question 3
3) Write a point form geological history of the picture above: Source: original question	**Relative Dating** – Apply knowledge of stratigraphic principles and sedimentary features to create a geological history	Analysis		-Replaced "geological history" with "sequence of events that would produce…" from students simply ordering units from the diagram instead of writing out events that involved the units -Made a specific reference to the diagram from question 2	

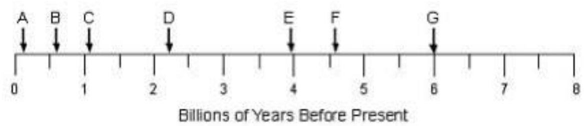
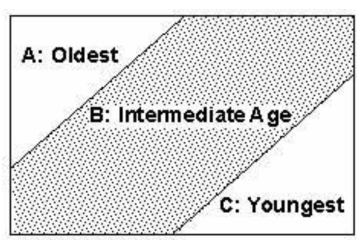
Appendix 1. Continued…

Changes in v4	Changes in v5	Changes in v6	Final Question
			1) Approximately how many years back in time did the Earth form? (a) 4 hundred years (b) 4 hundred-thousand years (c) 4 million years (d) 4 billion years (e) 4 trillion years
	-Definition of sides of erosion feature "X" in response to validation of question 3	-Bold and underline font on "best" to create consistency within test on key words	2) What is the **best** estimate of the age of F if A is 100 million years old and D is 70 million years old?  (a) 55 Myrs (b) 85 Myrs (c) 110 Myrs (d) 170 Myrs (e) same age as A (f) same age as D
			3) In point form, write a sequence of events that would produce the cross section from question 2:

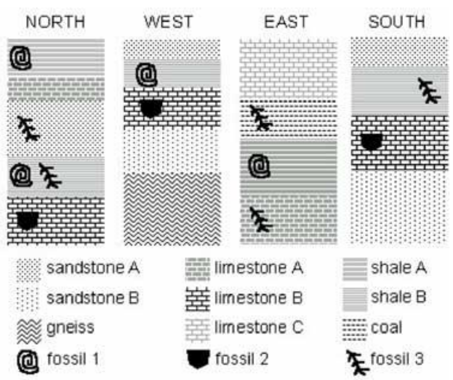
Appendix 1 continued

Draft of Question & Source	**Key concept from Table 2.**	**Bloom's Cognition Level**	**Changes in v1**	**Changes in v2**	**Changes in v3**
4) Carefully examine the relative positions of the lettered arrows on the timeline below and estimate the ages represented by each arrow. Identify which letter corresponds most closely to the extinction of the dinosaurs.  (a) A (b) B (c) C (d) D (e) E Source: http://serc.carleton.edu/6173	**Timescale** – Understand the immensity of geological time	Comprehension		-Replace distractors 'D' and 'E' with 'F' and 'G' to encompass entire range of diagram - from instructor input	-Bold and Underline font on "Carefully" to emphasize key word (Burton et al, 1991)
5) Which method was primarily used to establish the Geologic Time Scale? (a) Carbon-Nitrogen (b) Potassium-Argon (c) Relative dating (d) Rubidium-Strontium (e) Uranium-Lead. Source: http://www.usd.edu/esci/exams/geoltime.html	**Timescale** – Explain framework for its construction – that it was based upon succession of fossil types	Knowledge		-Completely replaced distractors because options (a), (b), (d), and (e) were too easily eliminated by students – knowing that the timescale is older than radiometric methods would immediately lead to the correct answer	
6) During fieldwork in the western U.S., an experienced geologist sketched the cross section below showing three different units of tilted rocks and their relative ages. What could you best infer from this diagram?  (a) Mountain building has overturned the units. (b) C formed first and so lies on the bottom. (c) B and C must be igneous rocks. (d) The sequence of rock layers conforms to the principle of superposition. (e) B is an intrusive unit Source: http://serc.carleton.edu/6138	**Relative Dating** – Apply relative dating principles to field and map interpretations to create geological histories	Application		-Bold and underline font on "best" because students were focusing on it during interviews -replaced (e) with "…original horizontality" because "…intrusive unit" was easily eliminated and to test knowledge of principles -added references to rock units to be as clear as possible since the diagram doesn't label rock types	-Changed "U.S." to "Canada" to increase relevancy to students at a Canadian institution – from instructor feedback

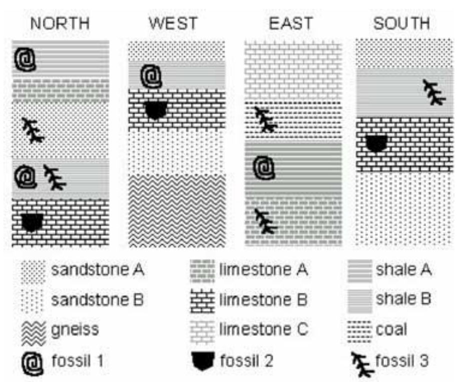
Appendix 1 continued…

Changes in v4	Changes in v5	Changes in v6	Final Question
			4) **<u>Carefully</u>** examine the relative positions of the lettered arrows on the timeline below and estimate the ages represented by each arrow. Identify which letter corresponds most closely to the extinction of the dinosaurs.  (a) A (b) B (c) C (d) F (e) G
	-Bold and underline font on "primarily" because students were reasoning for multiple distractors being correct instead of focusing on one correct answer	-Bold and underline font on "establish" for emphasis on key words (Burton et al, 1991) -Students were focusing on "establish" throughout validations	5) Which method was **<u>primarily</u>** used to **<u>establish</u>** the Geologic Time Scale? (a) Correlation of magnetic signatures in rocks (b) Calculation of alpha decay of isotopes (c) Calculation of beta decay of isotopes (d) Correlation of fossils in rock units across vast distances (e) Correlation of rock types across vast distances
			6) During fieldwork in western Canada, an experienced geologist sketched the cross section below showing three different units of tilted rocks and their relative ages. What could you **<u>best</u>** infer from this diagram?  (a) Mountain building has overturned the rock units. (b) Rock layer C formed first and so lies on the bottom. (c) Rock layers B and C must be igneous rocks. (d) The sequence of rock layers conforms to the principle of superposition. (e) The sequence of rock layers conforms to the principle of original horizontality.

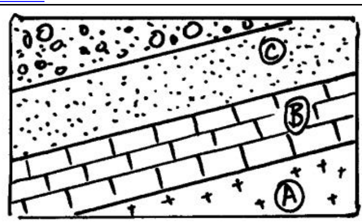
Appendix 1 continued

Draft of Question & Source	**Key concept from Table 2.**	**Bloom's Cognition Level**	**Changes in v1**	**Changes in v2**
7) Which answer best describes what the surface of the Earth would be like if you could travel back to the time when the Earth first formed as a planet? (a) The Earth's surface was covered with jungles (b) The Earth's surface was covered with water (e) The Earth's surface was similar to modern undeveloped areas (d) The Earth's surface was covered with melted rock (e) The Earth's surface was covered with ice Source: Geoscience Concept Inventory	**Earth History** – Describe the history of Earth's formation	Knowledge	-Removed reference to time travel – from instructor feedback	-Replaced distractors of "jungles", "modern undeveloped areas", and "ice" because they were too easily eliminated – from student validation and instructor feedback
8) Four outcrops of rock are examined in different locations of a state. The rock types and the fossils they contain are illustrated in the adjacent diagram. Which fossil would be the **best** choice to use as an index fossil for these rocks?  (a) fossil 1 (b) fossil 2 (c) fossil 3 Source: http://serc.carleton.edu/6168	**Relative Dating** – Know that fossils can be used to define units because they are separated in time	Application	-Addition of distractors (d)"There are no index fossils" and (e)" they can all be used as index fossils" to be consistent with other questions – from instructor feedback	-Replaced "a state" with "British Columbia" to increase relevancy at a Canadian institution -Replaced distractor (e) with "all fossils make equally good index fossils" for clear wording
9) The isotope Einsteinium-253 has a half-life of 20 days. If you began an experiment with an 80-gram sample of Einsteinium-253, how much would remain after 60 days? (a) 60 grams (b) 40 grams (c) 20 grams (d) 10 grams (e) not enough information provided Source: http://serc.carleton.edu/6159	**Absolute Dating** – Define the principles of radioactive decay	Analysis		

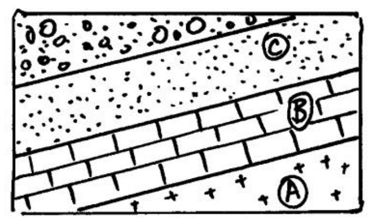
Appendix 1 continued…

Changes in v3	Changes in v4	Changes in v5	Changes in v6	Final Question
-Replace "formed" with "accreted" to define specific point in Earth's history – from discrepancy in students understanding of what "formed" referred to		-Bold and underline font on "best" to emphasize key words (Burton et al, 1991)		7) Which answer **best** describes what the surface of the Earth was like when the Earth first accreted as a planet? (a) The Earth's surface was covered with solid rock (b) The Earth's surface was covered with water (c) The Earth's surface was covered with solid continental masses (d) The Earth's surface was covered with melted rock (e) The Earth's surface was covered with melted iron
		-Add definition of index fossil: "An Index Fossil is a fossil that dates the strata in which it is found." – from students not knowing what an index fossil is during validations	-Underline font on "best" for consistency with other emphasized words	8) An Index Fossil is a fossil that dates the strata in which it is found. Four outcrops of rock are examined in different locations of British Columbia. The rock types and the fossils they contain are illustrated in the adjacent diagram. Which fossil would be the **best** choice to use as an index fossil for these rocks?  (a) fossil 1 (b) fossil 2 (c) fossil 3 (d) there are no index fossils (e) all fossils make equally good index fossils
				9) The isotope Einsteinium-253 has a half-life of 20 days. If you began an experiment with an 80-gram sample of Einsteinium-253, how much would remain after 60 days? (a) 60 grams (b) 40 grams (c) 20 grams (d) 10 grams (e) not enough information provided

Appendix 1 continued…

Draft of Question & Source	Key concept from Table 2.	Bloom's Cognition Level	Changes in v1	Changes in v2
10) What do we call the feature left by a cycle involving deposition, then removal of previously-deposited sediment by erosion, then a return to deposition? (a) Cross-cutting. (b) Inclusion. (c) Rock cycle. (d) Superposition. (e) Unconformity. Source: [http://www.usd.edu/esci/exams/geoltime.html #7](http://www.usd.edu/esci/exams/geoltime.html)	**Relative Dating** – Apply knowledge of stratigraphic principles and sedimentary features to create a geological history	Knowledge		-Replaced distractors of "Rock cycle" and "Superposition" with "Turbidite sequence" and "non conformity" to make all distractors features not processes – From instructor feedback -replacing distractors makes them more difficult for students to eliminate – from validations -Made distractors grammatically correct
 11) Layer A is from the Ordivician, Layer C is from the Carboniferous. What geological time Period is most likely represented by Layer B? (a) Paleozoic (b) Cretaceous (c) Devonian (d) Cambrian (e) Mesozoic Source: original question	**Time Scale** – Familiarity with terminology of geological eons, eras, and periods	Analysis		-Spelling of "Ordovician" corrected -Distractors (a), (b), (d), and (e) refined to reflect a more specific cluster along the timescale to be more difficult to eliminate – from instructor feedback due to false positive from student validation -Font bolded and underlined on "most likely" to provide emphasis
12) Mt. Everest, along with the rest of the Himalayan Plateau began forming… (a) In the Middle Proterozoic when India drifted Northward into Eurasia. (b) In the Late Cretaceous when Eurasia drifted Southward into India. (c) 70 Million years ago when India drifted Northward into Eurasia at a divergent boundary. (d) 70 Million years ago when India quickly moved Eastward into Eurasia. (e) In the Late Cretaceous when India drifted Northward into Eurasia. Source: original question	**Earth History** – Describe the paleo-geographical development of Earth & **Timescale** – Reproduce time scale to eras, periods, and epochs with the associated dates	Knowledge		

Appendix 1 continued…

Changes in v3	Changes in v4	Changes in v5	Changes in v6	Final Question
				10) What do we call the feature left by a cycle involving deposition, then removal of previously-deposited sediment by erosion, then a return to deposition? (a) A cross-cutting relationship (b) An inclusion (c) A turbidite sequence (d) A nonconformity (e) An unconformity
		-Font bolded and underlined on "Period" to provide emphasis on what the question is asking for (a specific named Period of the time scale not a portion of time)		 11) Layer A is from the Ordovician, Layer C is from the Carboniferous. What geological time **<u>Period</u>** is **<u>most likely</u>** represented by Layer B? (a) Lower Carboniferous (b) Upper Ordovician (c) Devonian (d) Permian (e) Paleozoic
			-Question was deleted because it was testing specific knowledge of tectonic history and timescale terminology – from difficulties in student validations	

Appendix 1 continued…

Draft of Question & Source	Key concept from Table 2.	Bloom's Cognition Level	Changes in v1	Changes in v2
13) Scientists have discovered fossils of four-legged creatures called dinosaurs. How much time passed between the appearance and extinction of these creatures? (a) Hundreds of years (b) Thousands of years (c) Millions of years (d) Billions of years (e) Some of these creatures still exist Source: Geoscience Concept Inventory	**Earth History** – Placement of major geological events on the timescale (extinctions, formations, plate movements etc…) with associated dates	Knowledge	-Removed preface of what dinosaurs are – unnecessary for geology majors – from instructor feedback -Replaced easily eliminated distractors (a), (b), and (e) with "Hundreds of Thousands", "Tens of Thousands", and "Hundreds of Billions"– from instructor feedback	-Replace "Hundreds of Billions" with "Hundreds of Millions" to make it harder to eliminate – from instructor feedback and student validation -Adjust distractors to be in numerical order
14) Some scientists claim that they can determine when the Earth first formed as a planet. Which technique(s) do scientists use today to determine when the Earth first formed? **<u>Choose all that apply.</u>** (a) Comparison of fossils found in rocks (b) Comparison of different layers found in rocks (c) Analysis of uranium and lead found in rocks (d) Analysis of carbon found in rocks (e) Scientists cannot calculate the age of the Earth Source: Geoscience Concept Inventory	**Absolute Dating** – Recognize commonly used radiometric dating methods	Comprehension	-Removed preface, "Some scientists claim that they can determine when the Earth first formed as a planet" – unnecessary for geology majors – from instructor feedback	-"Choose all that apply" replaced with "Choose the best answer." – to get intended selection of one correct answer – from student validation -(e)"Scientists cannot calculate…" replaced with "Correlation of fossils in different layers of rock" to remove negative answer (Burton et al, 1991) -(a) "Comparison of fossils…" replaced with "Cooling of metal spheres…" to provide historically correct distractor -(b) "…different layers found in rocks" replaced with "…different layers of rock" to be grammatically correct
15) An archaeologist is studying layers of human settlement in an African cave. What method of radiometric dating should they use to determine the age of woody material found at an apparent fire pit in one of the layers? (a) Potassium – Argon (b) Argon – Argon (c) Carbon – Nitrogen (d) Carbon – Helium (e) Not enough information is provided. Source: original question	**Absolute Dating** – Apply multiple dating processes appropriately in different geological settings	Application		-Distractor "Carbon – Helium" replaced with "Rubidium – Strontium" because Carbon – Helium was easily eliminated – from student validation -"Not enough information is provided" replaced by "Carbon – Oxygen" to remove negative distractor (Burton et al, 1991)

Appendix 1 continued…

Changes in v3	Changes in v4	Changes in v5	Changes in v6	Final Question
-Font bolded and underlined on "between" for emphasis -distractors changed from words to number ranges (ie. 100,000-1,000,000; 1,000,000-65,000,000; 65,000,000-100,000,000; 100,000,000 - 200,000,000; 200,000,000 - 500,000,000)		-time ranges replaced with specific lengths of time – From student validation and instructor feedback – there was a disconnect in asking students for a specific length of time and providing ranges as distractors	-Re-numbered to 12	12) How much time passed **between** the appearance and extinction of dinosaurs? (a) 35,000,000 years (b) 65,000,000 years (c) 100,000,000 years (d) 135,000,000 years (e) 200,000,000 years
	-"Which technique…" changed to "What technique…" for grammatical flow		-Re-numbered to 13	13) What technique do scientists use today to determine when the Earth first formed? Choose the **best** answer. (a) Calculation of the cooling rate of metal spheres (b) Comparison of different layers of rock (c) Analysis of uranium and lead found in rocks (d) Analysis of carbon found in rocks (e) Correlation of fossils in different layers of rock
			-Re-numbered to 14	14) An archaeologist is studying layers of human settlement in an African cave. What method of radiometric dating should they use to determine the age of woody material found at an apparent fire pit in one of the layers? (a) Potassium – Argon (b) Argon – Argon (c) Carbon – Nitrogen (d) Rubidium – Strontium (e) Carbon - Oxygen

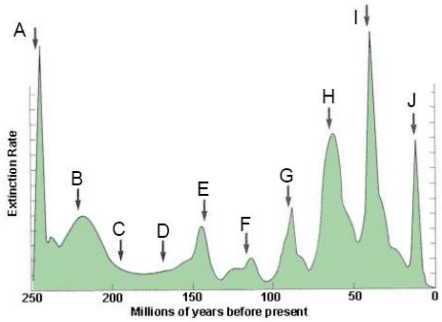
Appendix 1 continued…

Draft of Question & Source	Key concept from Table 2.	Bloom's Cognition Level	Changes in v1	Changes in v2
16) Why is U^{238}-Pb^{206} dating not the best method for determining the age of the woody debris from the previous question? Choose the BEST answer. (a) Uranium is only found in Canada, and the cave is located in Africa. (b) Fire in the pit would cause the woody debris to reach closure temperature and all daughter nuclides would be lost. (c) U^{238}-Pb^{206} dating would only reveal the time of formation of the woody debris, not the actual time of deposition. (d) U^{238}-Pb^{206} dating does not cover a sufficiently short time scale to be used on the sample. (e) Woody debris cannot be run through the analytical methods needed to determine U^{238}-Pb^{206} ratios Source: original question	**Absolute Dating** – Apply multiple dating processes appropriately in different geological settings	Application	-Bold font on "not" for emphasis on a key word(Burton et al, 1991)	-Distractor (d) "…does not cover a sufficiently short time scale to be used on the sample" replaced with "…is not accurate within the amount of time represented by the sample" to clarify misinterpretations during student validation
17) During the Triassic, the Tethys Ocean was representative of what paleo-geographic environment? (a) Cold climate with global ice coverage. (b) Temperate climate with deep ocean basins. (c) Warm climate with shallow seas. (d) Cold climate with shallow seas (e) The Tethys Ocean did not exist during the Triassic. Source: original question	**Earth History** – Describe the paleo-geographical development of Earth	Knowledge		
18) In the age equation, $\mathbf{D = D_0 + N(e^{\lambda t}\text{-}1)}$, the symbol λ represents: (a) alpha decay (b) radioactive half-life (c) radioactive decay constant (d) $(t^{1/2})$*(ln 2) (e) $(t^{1/2})$/(ln 2) Source: original question	**Rates and Processes** – Quantify geological processes from chemical and physical rate laws	Knowledge		-New question introduced

Appendix 1 continued…

Changes in v3	Changes in v4	Changes in v5	Changes in v6	Final Question
-"BEST" replaced with "**best**" for consistency and emphasis on a key word (Burton et al, 1991)		-Underline font on "not" for consistency and emphasis on a key word (Burton et al, 1991)	-Re-numbered to 15	15) Why is U^{238}-Pb^{206} dating **not** the best method for determining the age of the woody debris from the previous question? Choose the **best** answer. (a) Uranium is only found in Canada, and the cave is located in Africa. (b) Fire in the pit would cause the woody debris to reach closure temperature and all daughter nuclides would be lost. (c) U^{238}-Pb^{206} dating would only reveal the time of formation of the woody debris, not the actual time of deposition. (d) U^{238}-Pb^{206} dating is not accurate within the amount of time represented by the sample. (e) Woody debris cannot be run through the analytical methods needed to determine U^{238}-Pb^{206} ratios.
			-Re-numbered to 16	16) During the Triassic, the Tethys Ocean was representative of what paleo-geographic environment? (a) Cold climate with global ice coverage. (b) Temperate climate with deep ocean basins. (c) Warm climate with shallow seas. (d) Cold climate with shallow seas (e) The Tethys Ocean did not exist during the Triassic.
		-"lambda" added beside λ because students did not know the name of the symbol – from student validations	-Re-numbered to 17 -age equation described as the "general form" to clue students in that it's not the simplified half life equation – from student validations	17) In the general form of the age equation, $\mathbf{D = D_0 + N(e^{\lambda t}-1)}$, the symbol λ (lambda) represents: (a) alpha decay (b) radioactive half-life (c) radioactive decay constant (d) $(t^{1/2})*(\ln 2)$ (e) $(t^{1/2})/(\ln 2)$

Appendix 1 continued…

<table>
<tr><th>Draft of Question & Source</th><th>Key concept from Table 2.</th><th>Bloom's Cognition Level</th><th>Changes in v1</th><th>Changes in v2</th></tr>
<tr><td>19) Which of the following labeled events would correlate to the base of the Mesozoic and the base of the Cenozoic?


(a) A, G
(b) I, J
(c) A, H
(d) H, J
(e) E, I

Source: Dr. Stuart Sutherland, Department of Earth and Ocean Sciences, UBC</td><td>Earth History – Placement of major geological events on the timescale</td><td>Knowledge</td><td></td><td>-New question introduced
-</td></tr>
<tr><td>20) Place the following tectonic events into <u>chronological order</u> (oldest – youngest):

Formation of Pangaea (A) | Formation of the Himalayas (B) | Opening of the Atlantic Ocean (C) | Formation of the Appalachian Mountains (D)

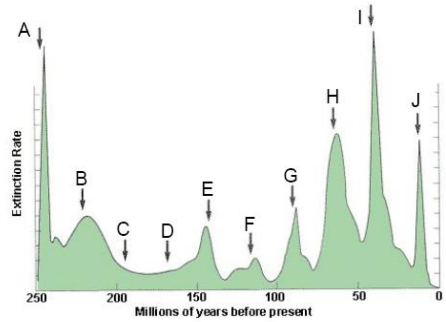
(a) A, B, C, D
(b) A, C, D, B
(c) B, D, A, C
(d) C, A, D, B
(e) D, A, C, B

Source: original question</td><td>Earth History – Know the relative timing of major geologic events</td><td>Comprehension</td><td></td><td></td></tr>
<tr><td><table><tr><th>A</th><th>B</th></tr><tr><td>Ripple Marks in Sedimentary beds
Trace fossils in mudstones
Exposed mountain roots</td><td>Streams feeding into lakes
Burrowing sea creatures
Chemical and physical erosion of mountains</td></tr></table>21) <u>Compare</u> columns A and B. What early-adopted geological principle can be used to explain the features and processes?

(a) Catastrophism
(b) Super position
(c) Uniformitarianism
(d) Gradualism
(e) Original horizontality

Source: original question</td><td>Uniformitariani sm – Know the concept and understand the context in which it was created</td><td>Analysis</td><td></td><td></td></tr>
</table>

Appendix 1 continued…

<table>
<tr><th>Changes in v3</th><th>Changes in v4</th><th>Changes in v5</th><th>Changes in v6</th><th>Final Question</th></tr>
<tr><td></td><td></td><td>-Font on “base” bolded and underlined for emphasis on key words – from student validations</td><td>-Renumbered to 18</td><td>18) Which of the following labeled events would correlate to the <u>base</u> of the Mesozoic and the <u>base</u> of the Cenozoic?


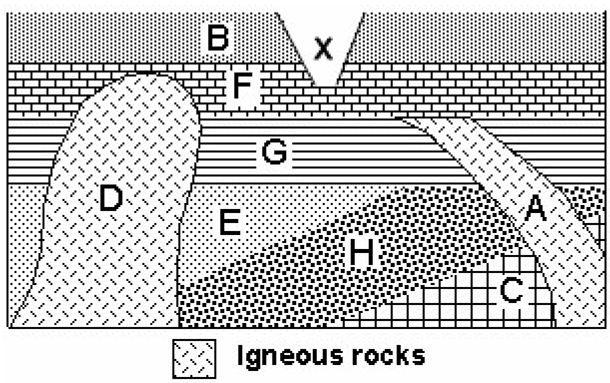
(a) A, G
(b) I, J
(c) A, H
(d) H, J
(e) E, I</td></tr>
<tr><td></td><td>-New question introduced</td><td></td><td>-Renumbered to 19</td><td>19) Place the following tectonic events into <u>chronological order</u> (oldest – youngest):
Formation of Pangaea (A) | Formation of the Himalayas (B) | Opening of the Atlantic Ocean (C) | Formation of the Appalachian Mountains (D)
(a) A, B, C, D
(b) A, C, D, B
(c) B, D, A, C
(d) C, A, D, B
(e) D, A, C, B</td></tr>
<tr><td></td><td>-New question introduced</td><td>-Table labeled with “Geological Features” and “Processes” to clarify the categorization – from student validations</td><td>-Renumbered to 20
-Distractors of “Superposition” and “Original Horizontality” replaced with “Lamarckism” and “Darwinism” to make distractors more appealing and consistent – Student Validation and Advice from Science Teaching and Learning Fellow
-Place question above table to ensure students read the question before making assumptions about the table – From student validations</td><td>20) <u>Compare</u> columns A and B. What early-adopted geological principle can be used to explain the features and processes?<table><tr><th>A</th><th>B</th></tr><tr><td>Ripple Marks in Sedimentary beds
Trace fossils in mudstones
Exposed mountain roots</td><td>Streams feeding into lakes
Burrowing sea creatures
Chemical and physical erosion of mountains</td></tr></table>(a) Catastrophism
(b) Lamarckism
(c) Uniformitarianism
(d) Gradualism
(e) Darwinism</td></tr>
</table>

Appendix 2. Final assessment mechanism and answer key.

1) Approximately how many years back in time did the Earth form?

(a) 4 hundred years
(b) 4 hundred-thousand years
(c) 4 million years
(d) 4 billion years
(e) 4 trillion years

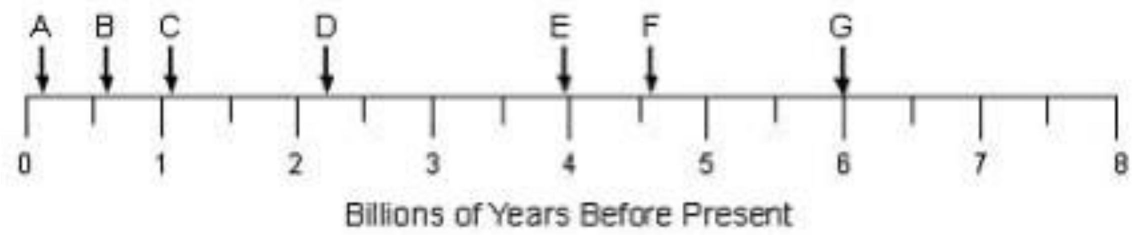
2) What is the **best** estimate of the age of F if A is 100 million years old and D is 70 million years old?



(a) 55 Myrs
(b) 85 Myrs
(c) 110 Myrs
(d) 170 Myrs
(e) same age as A
(f) same age as D

3) In point form, write a sequence of events that would produce the cross section from question 2:

4) **<u>Carefully</u>** examine the relative positions of the lettered arrows on the timeline below and estimate the ages represented by each arrow. Identify which letter corresponds most closely to the extinction of the dinosaurs.

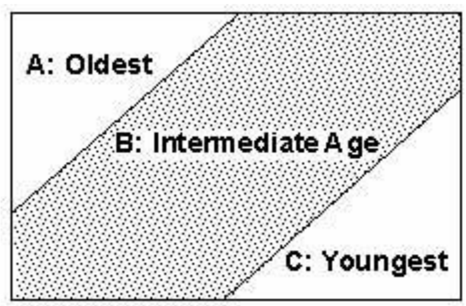


(a) A
(b) B
(c) C
(d) F
(e) G

5) Which method was **<u>primarily</u>** used to **<u>establish</u>** the Geologic Time Scale?

(a) Correlation of magnetic signatures in rocks
(b) Calculation of alpha decay of isotopes
(c) Calculation of beta decay of isotopes
(d) Correlation of fossils in rock units across vast distances
(e) Correlation of rock types across vast distances

6) During fieldwork in western Canada, an experienced geologist sketched the cross section below showing three different units of tilted rocks and their relative ages. What could you **best** infer from this diagram?



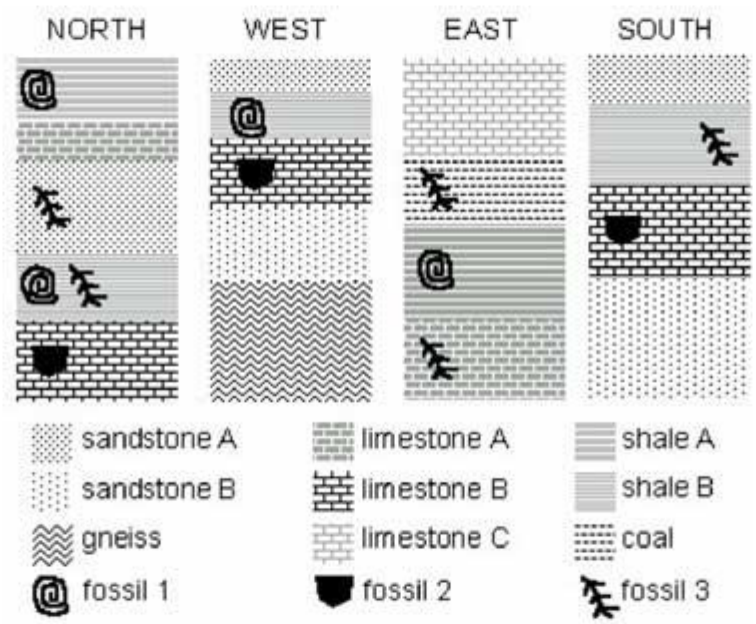
(a) Mountain building has overturned the rock units.
(b) Rock layer C formed first and so lies on the bottom.
(c) Rock layers B and C must be igneous rocks.
(d) The sequence of rock layers conforms to the principle of superposition.
(e) The sequence of rock layers conforms to the principle of original horizontality.

7) Which answer **best** describes what the surface of the Earth was like when the Earth first accreted as a planet?

(a) The Earth's surface was covered with solid rock
(b) The Earth's surface was covered with water
(c) The Earth's surface was covered with solid continental masses
(d) The Earth's surface was covered with melted rock
(e) The Earth's surface was covered with melted iron

8) An Index Fossil is a fossil that dates the strata in which it is found.

Four outcrops of rock are examined in different locations of British Columbia. The rock types and the fossils they contain are illustrated in the adjacent diagram. Which fossil would be the **best** choice to use as an index fossil for these rocks?



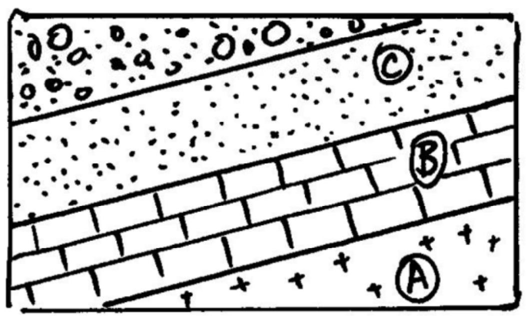
(a) fossil 1
(b) fossil 2
(c) fossil 3
(d) there are no index fossils
(e) all fossils make equally good index fossils

9) The isotope Einsteinium-253 has a half-life of 20 days. If you began an experiment with an 80-gram sample of Einsteinium-253, how much would remain after 60 days?

(a) 60 grams
(b) 40 grams
(c) 20 grams
(d) 10 grams
(e) not enough information provided

10) What do we call the feature left by a cycle involving deposition, then removal of previously-deposited sediment by erosion, then a return to deposition?

(a) A cross-cutting relationship.
(b) An inclusion.
(c) A turbidite sequence.
(d) A nonconformity.
(e) An unconformity.



11) Layer A is from the Ordovician, Layer C is from the Carboniferous. What geological time **<u>Period</u>** is **<u>most likely</u>** represented by Layer B?

(a) Lower Carboniferous
(b) Upper Ordovician
(c) Devonian
(d) Permian
(e) Paleozoic

12) How much time passed **<u>between</u>** the appearance and extinction of dinosaurs?

(a) 35,000,000 years
(b) 65,000,000 years
(c) 100,000,000 years
(d) 135,000,000 years
(e) 200,000,000 years

13) What technique do scientists use today to determine when the Earth first formed? Choose the **best** answer.

(a) Calculation of the cooling rate of metal spheres
(b) Comparison of different layers of rock
(c) Analysis of uranium and lead found in rocks
(d) Analysis of carbon found in rocks
(e) Correlation of fossils in different layers of rock

14) An archaeologist is studying layers of human settlement in an African cave. What method of radiometric dating should they use to determine the age of woody material found at an apparent fire pit in one of the layers?

(a) Potassium – Argon
(b) Argon – Argon
(c) Carbon – Nitrogen
(d) Rubidium – Strontium
(e) Carbon - Oxygen

15) Why is U^{238}-Pb^{206} dating **not** the best method for determining the age of the woody debris from the previous question? Choose the **best** answer.

(a) Uranium is only found in Canada, and the cave is located in Africa.
(b) Fire in the pit would cause the woody debris to reach closure temperature and all daughter nuclides would be lost.
(c) U^{238}-Pb^{206} dating would only reveal the time of formation of the woody debris, not the actual time of deposition.
(d) U^{238}-Pb^{206} dating is not accurate within the amount of time represented by the sample.
(e) Woody debris cannot be run through the analytical methods needed to determine U^{238}-Pb^{206} ratios.

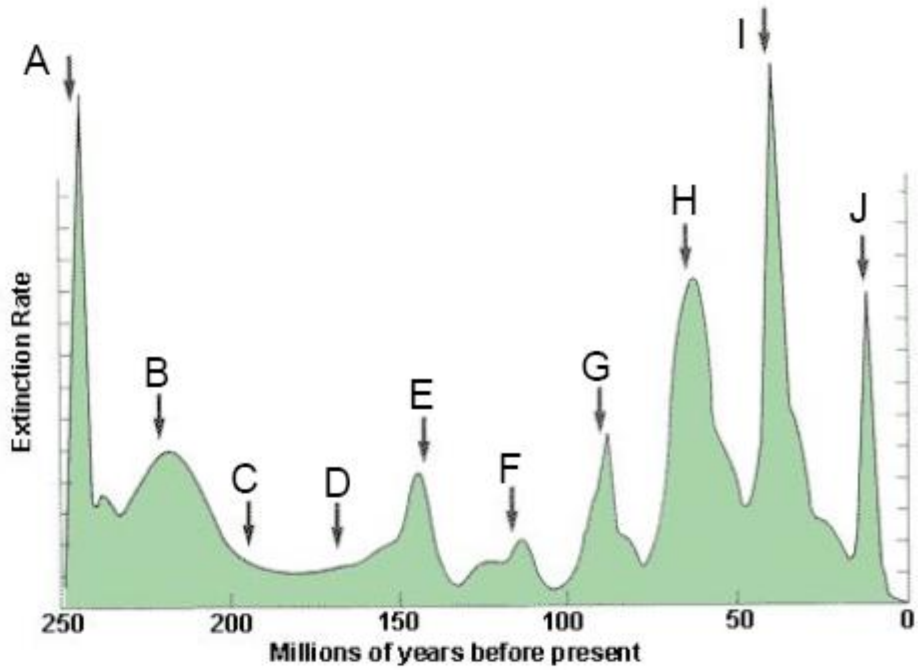
16) During the Triassic, the Tethys Ocean was representative of what paleo-geographic environment?

(a) Cold climate with global ice coverage.
(b) Temperate climate with deep ocean basins.
(c) Warm climate with shallow seas.
(d) Cold climate with shallow seas
(e) The Tethys Ocean did not exist during the Triassic.

17) In the general form of the age equation, $\mathbf{D = D_0 + N(e^{\lambda t}-1)}$, the symbol λ (lambda) represents:

(a) alpha decay
(b) radioactive half-life
(c) radioactive decay constant
(d) $(t^{1/2})*(\ln 2)$
(e) $(t^{1/2})/(\ln 2)$

18) Which of the following labeled events would correlate to the **base** of the Mesozoic and the **base** of the Cenozoic?



(a) A, G
(b) I, J
(c) A, H
(d) H, J
(e) E, I

19) Place the following tectonic events into **chronological order** (oldest – youngest):

Formation of Pangaea	Formation of the Himalayas	Opening of the Atlantic	Formation of the Appalachian

(a) A, B, C, D
(b) A, C, D, B
(c) B, D, A, C
(d) C, A, D, B
(e) D, A, C, B

20) **<u>Compare</u>** columns A and B. What early-adopted geological principle can be used to explain the features and processes?

A – Geological Features	**B - Processes**
Ripple Marks in Sedimentary beds	Streams feeding into lakes
Trace fossils in mudstones	Burrowing sea creatures
Exposed mountain roots	Chemical and physical erosion of mountains

(a) Catastrophism
(b) Lamarckism
(c) Uniformitarianism
(d) Gradualism
(e) Darwinism

Answer Key:

Question	Answer
1	(d)
2	(b)

3	▪ Layer C deposited, Layer H deposited, Layer E deposited
	▪ Layers C,H,E tilted and eroded
	▪ Layer G deposited
	▪ Unit A intrudes layers C,H, and G
	▪ Unit A eroded
	▪ Layer F deposited
	▪ Order of these options can be reversed or listed as contemporaneous:
	    ▪ Unit D intrudes layers E,H, G, and F or
	    ▪ Layer B deposited
	▪ Layers B and F eroded, resulting in feature X

4	(a)
5	(d)
6	(a)
7	(d)
8	(b)
9	(d)
10	(e)
11	(c)
12	(d)
13	(c)
14	(c)
15	(d)
16	(c)
17	(c)
18	(c)
19	(e)
20	(c)

Appendix 3. Ethics consent form for validation interview participants.

THE UNIVERSITY OF BRITISH COLUMBIA

Carl Wieman Science Education Initiative (CWSEI)
University of British Columbia
Wesbrook Bldg.
300-6174 University Blvd.
Vancouver, BC Canada V6T 1Z3

Tel: (604) 827-3119
Fax: (604) 827-3118

Consent Form #2

Development and Validation of Surveys Probing Students' Understanding of Concepts in Science

Principal Investigator: Dr. Carl Wieman
Carl Wieman Science Education Initiative (CWSEI)
604-822-1732

Co-Investigator: Dr. Sarah Gilbert
Carl Wieman Science Education Initiative (CWSEI)
604-822-3193

Study Team Members/Researchers:

Jackie Stewart, Chemistry, 604-822-5912
Jennifer Duis, Chemistry, 604-827-5969
Steven Wolfman, Computer Science, 604-822-0407
Sara Harris, Earth and Ocean Sciences, 604-822-9651
Tom-Pierre Frappe-Seneclauze, Earth and Ocean Sciences, 604-822-3063
Brett Gilley, Earth and Ocean Sciences, 604-822-2138
Francis Jones, Earth and Ocean Sciences, 604-822-2138
Erin Lane, Earth and Ocean Sciences, 604-822-2138
Bruce Dunham, Statistics, 604-822-4997
Eugenia Yu, Statistics, 604-827-5756
Nancy Heckman, Statistics, 604-822-3595
Leah MacFadyen, Life Sciences, 604-827-3001
George Spiegelman, Life Sciences, 604-822-2036
Joanne Nakonechny, Life Sciences, 604-822-4691
Gulnur Birol, Life Sciences, 604-827-3414
Kathy Nomme, Life Sciences, 604-822-4788
Carol Pollock, Life Sciences, 604-822-4984

Ellen Rosenberg, Life Sciences, 604-822-6584
Harald Yurk, Life Sciences, 604-822-2582
Jared Taylor, Life Sciences, 604-822-2582
Jim Carolan, Physics & Astronomy, 604-822-3853
James Day, Physics & Astronomy, 604-822-1997
Peter Newbury, Physics & Astronomy, newbury@phas.ubc.ca
Matthew Martinuk, Physics & Astronomy, Martinuk@physics.ubc.ca
Louis Deslauriers, Physics & Astronomy, louisd@phas.ubc.ca
Costanza Piccolo, Mathematics, costanza@math.ubc.ca
Maura MacPhee, Nursing, maura.macphee@nursing.ubc.ca
Benjamin Yu, Computer Science, benyu@cs.ubc.ca
Michael Carlson, Earth & Ocean Science, carlsom@interchange.ubc.ca
Joshua Caulkins, Earth and Ocean Science, jcaulkins@eos.ubc.ca
Rebekah Mohr, Statistics, rebekah@stat.ubc.ca
Yuri Samozvanov, Chemistry, yuri_s@interchange.ubc.ca
Jamil Rhajiak, Earth and Ocean Science, jrhajiak@gmail.com

UBC undergraduate students are invited to participate in a study aimed at improving the learning and appreciation of science. This study is conducted by science departments at UBC and the Carl Wieman Science Education Initiative (CWSEI). Conclusions and the survey instruments based on composite interviews of many students may be published in some form and/or presented publicly, but without any information that could be used to identify the participants.

Purpose:
The purpose of this study is to develop and validate surveys that examine university students' understanding of concepts in particular science disciplines. This study will test students' understanding of particular topics and survey questions on those topics through interviews with a range of students that span the populations for which the surveys will be used. Interviews are an important development and "validation" step in the development of a survey; they explore how students think and learn a topic and test whether survey questions are worded so that the students interpret them as intended.

Study Procedures:
Your participation will involve watching online learning materials, listening to questions, or filling out a survey on some key concepts in a science discipline. Then the interviewer will ask questions in order to understand how you think about the topic, interpret each question, and the reasons for your particular responses to the survey question or to questions or online materials. They will proceed in this fashion through the survey, or, set of online materials, or conceptual topics to be covered in this interview. Typical interviews will last no more than one hour. All interviews will be audio recorded and may be video recorded with your permission.

Participants will be selected based on their grades to ensure an equal representation of students from each grade band. In the event that there are more participants than needed, subjects may be selected based on gender, ethnic background, and age for the sole purpose of ensuring the best match to the demographic representation of the overall UBC student population. At the time of recruitment, researchers will ask students for their ID numbers and permission to review their grades and background information for this purpose of participant selection.

Potential Risks:
There are no known risks to participation in this study.

Potential Benefits:
The benefits to you are indirect; the surveys developed as a result of these interviews are part of a major UBC initiative to improve science education. Input from the surveys is an essential component in understanding what changes in educational approaches are working well and where further improvements are needed. This may result in improvements to science courses you take in future semesters.

The benefits to society in general will be improved science education that most students will find more interesting and relevant to their lives.

Confidentiality:
Your confidentiality will be respected. Interviews will be transcribed and no one except the researchers will have access to your identity. The interviewer will be a faculty or staff researcher from each participating department with expertise in the respective discipline and will not be an instructor of any course in which you are currently enrolled. Any written or printed out materials with identifiable information will be stored in a locked filing cabinet and will not be available to any of your current instructors. Any information in electronic format will be stored on password protected computers. No individual student identifiers will be used in any published or publicly presented work.

Remuneration/Compensation:
For your participation, you will receive the monetary compensation of $15.

Contact for information about the study:
If you have any questions or would like further information about this study, you may contact the appropriate researcher for the particular survey instrument as follows:

Chemistry – Jackie Stewart (604-822-5912)
Jennifer Duis (604-827-5969)
Yuri Samozvanov, yuri_s@interchange.ubc.ca

Computer Sciences – Steven Wolfman (604-822-0407)
Benjamin Yu (604-827-3005)

Earth and Ocean Sciences – Sara Harris (604-822-9651)
Tom-Pierre Frappe-Seneclauze (604-822-3063)
Brett Gilley (604-822-2138)
Francis Jones (604-822-2138)
Erin Lane (604-822-2138)
Michael Carlson (carlsom@interchange.ubc.ca)
Joshua Caulkins (604-822-2138)
Jamil Rhajiak (jrhajiak@gmail.com)

Statistics – Bruce Dunham (604-822-4997)
Eugenia Yu (604-827-5767)
Nancy Heckman (604-822-3595)
Rebekah Mohr (rebekah@stat.ubc.ca)

Life Sciences – Leah MacFadyen (604-827-3001)
George Spiegelman (604-822-2036)
Joanne Nakonechny (604-822-4691)
Gulnur Birol (604-827-3414)
Kathy Nomme (604-822-4788)
Carol Pollock (604-822-4984)
Ellen Rosenberg (604-822-6584)
Harald Yurk (yurk@zoology.ubc.ca)
Jared Taylor (jtaylor@zoology.ubc.ca)

Physics & Astronomy - Jim Carolan (604-822-3853)
James Day (604-822-1997)
Peter Newbury, newbury@phas.ubc.ca
Matthew Martinuk, Martinuk@physics.ubc.ca
Louis Deslauriers, louisd@phas.ubc.ca

Mathematics – Costanza Piccolo (604- 827-3299)

Nursing – Maura MacPhee (604- 822-2891)

Contact for concerns about the rights of research subjects:
If you have any concerns about your treatment or rights as a research subject, you may contact the Research Subject Information Line in the UBC Office of Research Services at 604-822-8598 or if long distance e-mail to RSIL@ors.ubc.ca.

Consent:
Your participation in this study is entirely voluntary and you may refuse to participate or withdraw from the study at any time without jeopardy to your class standing.

*If you require additional time to review this consent form, please feel free to do so and return the signed form to the appropriate researcher by tomorrow.

Your signature below indicates that you have received a copy of this consent form for your own records.

Your signature indicates that you consent to participate in this study.

____

Participant's Signature Date

____

Printed Name of the Participant signing above

Appendix 4. Certificate of completion of the TCPS: Ethical Conduct for Research Involving Humans

Certificate of Completion

This is to certify that

Jamil A. N. Rhajiak

has completed the Interagency Advisory Panel on Research Ethics' Introductory Tutorial for the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS)

Issued On: February-2-2009