On The Validity Of Concept Map-Base Assessment Interpretations: An Experiment Testing The Assumption Of Hierarchical Concept Maps In Science

CSE Technical Report 455

Maria Araceli Ruiz-Primo
Richard J. Shavelson
Susan Elise Schultz
Stanford University

July, 1997

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Center for the Study of Evaluation (CSE)
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, Ca 90095-6511
(310) 206-1532

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

ON THE VALIDITY OF CONCEPT MAP-BASE ASSESSMENT INTERPRETATIONS: AN EXPERIMENT TESTING THE ASSUMPTION OF HIERARCHICAL CONCEPT MAPS IN SCIENCE

Maria Araceli Ruiz-Primo
Richard J. Shavelson
Susan Elise Schultz
Stanford University

Abstract

A concept map consists of a *task*, a *response format*, and a *scoring system*. Variation in tasks, response formats, and scoring systems may elicit different knowledge representations, posing construct-interpretation challenges. This study examined two mapping techniques: Technique 1 asked students to construct a hierarchical map and Technique 2 did not impose any structure. Regardless of organization, we expected that as subject-matter knowledge increases, the structure of the map should increasingly reflect the structure of the domain as held by experts. Two types of topics were used, one hierarchically (Atom Structure) and the other non-hierarchically (Ions, Molecules and Compounds) structured. Topics were selected as having different structures according to experts' concepts maps. Topic and Mapping Techniques were the two factors in the 2x2 factorial design used in this study. Three types of map scores were used: proposition accuracy—sum of scores obtained on all propositions; convergence—the proportion of valid propositions in the student's map out of all propositions in the criterion map; and salience—the proportion of valid propositions out of all the propositions in the students' map. Preliminary results indicate high interrater reliability coefficients across the types of scores. No significant interaction effect, topic by mapping technique, was found in any type of score.

Introduction

As expertise in a domain grows, through learning, training, and/or experience, the elements of knowledge become increasingly interconnected (e.g., Glaser & Bassok, 1989; Shavelson, 1972). To be knowledgeable in a domain implies a highly integrated conceptual structure assuming that knowledge within a content domain is organized around central concepts. Concept interrelatedness, then, is an essential property of knowledge. Indeed, one characteristic of expertise in a domain is a highly integrated knowledge structure.

Concept maps, proposed as a supplement to traditional multiple-choice tests for classroom and even large-scale assessment use (e.g., Lomask, Baron, Greig, & Harrison, 1992; Barenholz & Tamir, 1992) are purported to measure the structure of a student's declarative knowledge.[1] The rationale behind this claim is that "the essence of knowledge is structure" (Anderson, 1984, p.5) and this structure may be captured with graphical/structural representations (e.g., Goldsmith, Johnson, & Acton, 1991; Jonassen, Beissner, & Yacci, 1993; White & Gunstone, 1992).

A concept map is a graphical representation consisting of nodes and labeled lines. The nodes correspond to important terms (standing for *concepts*) in a domain.[2] The lines denote a relation between a pair of concepts (nodes). And the label on the line tells how the two concepts are related. The combination of two nodes and a labeled line is called a *proposition*—the basic unit of meaning in a concept map and the smallest unit that can be used to judge the validity of the relationship drawn between two concepts (e.g., Dochy, 1996).

Before concept maps are used in classrooms or for large-scale assessment and map scores are reported to teachers, students, the public, and policy-makers, research needs to provide information about their technical characteristics. Over the past three years, we have done research intended to create and inform a concept-map-assessment knowledge base. Our goals have been to provide not

---

[1] The term "assessment" reflects our belief that reaching a judgment about an individual's knowledge and skills requires an integration of several pieces of information; we consider concept maps as potentially one of those pieces (see Cronbach, 1990).

[2] Actually, terms or words used in concept mapping are not concepts. They stand for concepts. Nevertheless, the terms used in concept mapping are called "concepts" and from here on out, we will follow this convention.

only evidence about reliability and validity of concept map assessments, but also a framework that can guide others' research in this area (e.g., Ruiz-Primo & Shavelson, 1996; Ruiz-Primo, Schultz, & Shavelson, 1996; Ruiz-Primo & Shavelson, 1997). Accordingly, the study reported here provides evidence bearing on the reliability and validity of concept maps as representations of students' knowledge structure. We examine whether traditional instructions to construct a hierarchical map are necessary, considering that the map should reflect the structure of the subject domain as represented in a student's memory rather than a preconceived psychological theory.

A Concept Map-Based Measurement Framework

We have conceived of a concept-map-based assessment as composed of: (a) a *task* that invites students to provide evidence bearing on their knowledge structure in a content domain; (b) a format for the student's *response*, and (c) a *scoring system* by which the student's concept map can be evaluated accurately and consistently. Without these three components, a concept map cannot be considered to be an assessment. By taking into account all possible tasks, response formats, and scoring system options reported in the literature, our characterization has made evident the enormity of variations in concept mapping techniques used in research and practice (e.g., Ruiz-Primo & Shavelson, 1996).

We identified different ways in which concept map tasks, response formats, and scoring systems varied (see Table 1). Variations among concept map tasks are: (a) *task demands*—the demands made on the students in generating their concept maps (e.g., students can be asked to fill-in a skeleton map, or construct a map from scratch, or talk about the relation between concepts); (b) *task constraints*—the restrictiveness of the task (e.g., students may or may not be asked to construct a hierarchical map, or to use one or more links between concepts, or to provide the concepts for the map); and (c) *task content structure*—the intersection of the task demands and the constraints of the structure of the subject-domain to be mapped (e.g., there is no need to impose a hierarchical structure if the content structure is not hierarchical).

Table 1

Concept Map Components and Variations Identified.

Map Assessment Components	Variations	Instances
TASK	• Task Demands	Students can be asked to: • fill-in a map • construct a map from scratch • organize cards • rate relatedness of concept pairs • write an essay • respond to an interview
	• Task Constraints	Students may or may not be: • asked to construct a hierarchical map • provided with the concepts used in the task • provided with the concept links used in the task • allowed to use more than one link between nodes • allowed to physically move the concepts around until a satisfactory structure is arrived at • asked to define the terms used in the map • required to justify their responses • required to construct the map collectively
	• Content Structure	The intersection of the task demands and constraints with the structure of the subject domain to be mapped.
RESPONSE	• Response Mode	Whether the student response is: • paper-and-pencil • oral • on a computer
	• Format Characteristics	Format should fit the specifics of the task
	• Mapper	Whether the map is drawn by a: • student • teacher or researcher
SCORING SYSTEM	• Score Components of the Map	Focus is on three components or variations of them: • propositions • hierarchy levels • examples
	• Use of a Criterion Map	Compare a student's map with an expert's map. Criterion maps can be obtained from: • one or more experts in the field • one or more teachers • one or more top students
	• Combination of Map Components and a Criterion Map	The two previous strategies are combined to score the student's maps.

Three types of response variation were identified in concept mapping: (a) the *response_mode*—whether the student's response is paper-and-pencil, oral, on a computer (e.g., students may be asked to draw the concept map on a piece of paper or to enter the concepts and the relations in a computer); (b) *response format*—the characteristics of the response depending upon the task, usually fitting the specifics of the task (e.g., if the task asks students to fill in the skeleton map, the skeleton map and the concepts will be provided to the students); and (c) *the_mapper*—who draws the map (e.g., students, teachers, interviewer).

We found three scoring strategies: (a) *score map components* (e.g., the number of nodes, links, cross-links); (b) *compare a student's map with a criterion map* (e.g., expert's concept map); and (c) *a combination of both strategies* (e.g., expert's concept maps are used to validate student's links and concepts).

If one only combines each of the 6 task demands with each of the 8 types of task constraints, there are hundreds of ways to produce a concept map (i.e., $6 \times 2^8 - 1$). Table 2 presents some examples of different types of tasks, response formats, and scoring systems used in practice and in research on concept maps (see Ruiz-Primo & Shavelson, 1996, for more examples). From these examples, it is clear that concept mapping techniques can vary widely in the way they elicit a student's knowledge structure, which in turn can produce different representations and scores. Still, all mapping techniques are assumed to measure the same construct, some aspect of a student's knowledge structure. If concept maps are to be used as a measurement tool, we must take the time and effort to provide evidence on the impact of different mapping techniques for representing a student's knowledge structure.

Three questions have guided our research in pursuing this goal: (1) Do different mapping techniques provide the same information about a student's knowledge structure? We suspect that different mapping techniques tap different aspects of cognitive structure and lead students to produce different concept maps. (2) Can raters reliably score concept maps? We wanted to see whether high interrater reliability can be obtained when different scoring criteria are used than those reported in previous studies (e.g., counting number of nodes). To this end, we created a scoring system that could capture the quality of the propositions in the students concept maps. (3) Do concept map-based assessments provide information about students' knowledge similar to that provided by a multiple-choice test? If so, maybe nothing new can be learned from concept maps.

Table 2
Five examples of Different Types of Tasks, Response Format and Scoring Systems Used in Research of Concept Maps

Authors	Task	Response	Scoring System
• Anderson & Huang (1989)	Fill-in a map on types of muscles and their functions using the 15 concepts and 6 linkages terms provided.	Paper-an-pencil response. Students filled in a prestructured skeleton map.	Combination of scoring a student's map components and comparison with a criterion map. Students' map propositions were classified into 20 accuracy categories according to the criterion map.
• Fisher (1990)	Task 1. Enter concepts and relation names in the computer with as many links as desired. Task 2. Fill-in-the-blank when a central concept is masked and the other nodes are provided.	Computer response in both tasks. Students construct their maps on a blank screen for task 1, and filled-in the node(s) in a skeleton map for task 2.	The author only proposed the SemNet computer program as an assessment tool, but did not present any scoring system to evaluate the maps.
• Lomask, Baron, Greig, Harrison (1992)	Write an essay on two central topics on biology (i.e., growing plant and blood transfusion).	Paper-and-pencil response. Trained teachers construct a map from students' written essay. No effort was made to elicit any hierarchy.	Comparison with a criterion map. Two structural dimensions were identified for the comparison: the *size* and the *strength* of structure. The final scored was based on the combination of both dimensions.
• Markham , Mintzes, & Jones (1994)	Construct a hierarchical concept map from 10 given concepts on mammals.	Paper-and-pencil response. Students drew the concept map on a blank page.	Score based on map components: number of concepts, relations, branching, hierarchies, crosslinks, and examples. Number of concepts and relations were taken as indications of the extent of students' knowledge.
• Nakhleh & Krajcik (1991)	Semi-structured interview about acids and bases.	Oral response. The interviewer drew three concepts maps—one for acids, one for bases, and one for pH—based on statements that revealed the student's propositional knowledge.	Score based on map components: Propositions and examples, cross-links, hierarchy. Experts' maps were used to identify critical nodes and relationships.

• Wallace & Mintzes (1990)	Construct a hierarchical concept map from ten given concepts on life zones.	Paper-and-pencil response. Students drew the concept map on a blank page.	Score based on map components: number of relationships, levels of hierarchy, branchings, cross-links, and general-to-specific examples.

Deciding which techniques should be preferred or studied is not an easy task. Unfortunately, cognitive theory does not provide an adequate basis to make a decision. Furthermore, current cognitive theories may be limited in their ability to guide mapping techniques because they tend to be middle-range theories focused on particular aspects of cognition. In our research we have applied the following criteria to narrow down alternatives: (a) appropriateness of the cognitive demands required by the task; (b) appropriateness of a structural representation in a content domain; (c) appropriateness of the scoring system used to evaluate the accuracy of the representation; and (d) practicality of the technique. We eliminated, for example, a fill-in-the-blank task because we regarded it as inappropriate for measuring students' knowledge structures—the task itself too severely restricted the students' representations. We also favored scoring criteria that focused on the adequacy of propositions. Finally, since our focus is on large-scale assessment, we eliminated mapping techniques that required one-to-one interaction between student and tester on practical grounds.

The techniques used in our research have been selected from the same task demand, "construct a map." Within this demand, task constraints were varied in different studies. Response format (viz. draw a map on a piece of paper) and scoring system (viz. scoring based on a criterion map and the quality of the propositions) were held constant across studies.

In an initial study we (Ruiz-Primo & Shavelson, 1997) examined the effect of providing/not providing concepts for constructing the map by varying the source of concept sample: students or assessor. Mapping Technique 1 asked students to provide 10 concepts in a domain with which to construct the map and Mapping Technique 2 provided 10 concepts. To test whether concept map scores were sensitive to the provision of concepts, two different concept samples, A and B, from the same domain were used in Technique 2. All students that participated in the study completed a multiple-choice test, then constructed a map with Technique 1, and finally constructed two maps with Technique 2. After constructing the map with Technique 1, half of the students used Sample A first and then Sample B (both lists have four concepts in common), and the other half used the samples in the opposite order. Results from this study indicated that both techniques provided scores with similar means, variances, and reliability and validity coefficients. Moreover, students' map scores did not differ, on average, from Sample A to Sample B since no significant differences in means or variances were found across samples of concepts.

We intend to continue building a concept-map-assessment knowledge base by comparing another two mapping techniques. In this study we examined a common practice when using concept maps: to ask students to construct hierarchical maps (e.g., Novak, Gowin, & Johansen, 1983; Markham, Mintzes, & Jones, 1994; Roth & Roychoudhury, 1993).

Hierarchical Structures in Concept Maps

No attention has been directed to how the mapping instructions interact with the structure of the subject domain to be mapped. This interaction is the focus of this study. Methodologically and conceptually, there is no need to impose a hierarchical structure on concept maps if the structure of the content domain to be mapped is not hierarchical. In fact, it may be that different map structures are needed to represent different types of content structures. For example, Harnisch, Sato, Zheng, Yamaji, and Connell (in press) proposed the use of "chain maps" to represent procedural or sequential activities. Regardless of the type of organization, we expect that as subject matter mastery increases, the structure of the map should increasingly reflect the structure, hierarchical or not, in the domain as held by experts.

For identifying the structure of a domain, we need to assume that there is some "ideal organization" that best reflects the structure, and that "experts" in that domain possess that ideal organization to some degree. Experts' knowledge structures are assumed to be highly connected and articulated (e.g., Glaser, in press). But, do all experts in a field share the same knowledge structure? Acton, Johnson, and Goldsmith (1994) showed that experts' structures are highly variable. Indeed, individual differences in experts' maps will arise because knowledge structure should reflect not only domain knowledge, but also a personal schema for thinking and cognitive activity (e.g., strategies for problem solving and interpretation; Glaser, in press). Therefore, we expected different experts to provide somewhat different concept maps and, consequently, inferences about the structure of a subject domain from one expert's knowledge structure to another might also vary.[3]

Assuming that any expert's knowledge structure provides a reasonable representation of the subject domain, how can we determine whether the structure is hierarchical? The identification of hierarchical structures from the natural (i.e., inorganic and organic), conceptual, and artifactual worlds (e.g., computer language, social events) has been a topic of discussion for the last three decades (e.g., Whyte, Wilson, & Wilson, 1969; Dress & von Haeseler, 1990). Unfortunately, the term hierarchy has been considered a "catch-all" term used to cover a variety of related yet distinct notions (e.g., Bunge, 1969; Green, 1969; Mesarovic & Macko, 1969). This makes it difficult to find a formal definition that can be used without controversy (e.g., Dress & von Haeseler, 1990; Green, 1969; Jones, 1969; Rosen, 1969).

Bunge, in 1969, proposed a formal definition of hierarchy: "Strictly speaking, a hierarchy or hierarchical structure is a set equipped with a relation of domination or its converse, subordination" (p. 17). According to his definition, H (i.e., a set of elements with binary relations) is a hierarchy *if and only if*: (1) H has one and only one beginner element—"a supreme commander"; (2) no matter how low in the hierarchy an element is, it is under the command of the beginner; (3) every member has a single boss; (4) the relation among the elements is asymmetric and transitive (in Bunge's colloquial terms, "Togetherness but no back talking;" p. 16); and (5) the relation between elements is a relation of domination or power (i.e., elements are held together by a subordinate relation). According to Bunge, any structure has to meet each of the five assumptions if it is to qualify as a hierarchy. In sum, "A diagram of a hierarchy is a finite tree branching out of a single point (namely *b*) and no loops" (Bunge, 1969, p. 19).[4]

---

[3] An entire study could be carried out on similarity of knowledge structures among experts.

[4] If a structure is hierarchical, another characteristic emerges: *levels*. Level is an ambiguous term that is also the object of philosophical debate (e.g., Bunge, 1969; Mesarovic, & Macko, 1969). A level can be considered as an "assembly of things of a defined kind, e.g., collection of systems characterized by a definite set of properties and laws..." (Bunge, 1969, p. 20). To define a hierarchy level, then, we should, for example, evaluate whether every member at a certain level shares an exclusive property that makes that level different from another level.

According to this definition "pure" hierarchical concept-map structures may be difficult to find: maps constructed by experts or knowledgeable students may not comply with criteria 2, 3, 4 and 5 since highly connected structures with crosslinks across levels and between branches are typical of mastered knowledge. Therefore, "degree of hierarchiness," may be a more accurate way to describe concept-map structures. A concept map that has more than one beginner node, many nodes with more than "one boss," many cycles (or loops), and concepts that are not held together by subordinate relations exclusively, can be considered "less hierarchical than" a map that has one beginner node, no nodes with more than "one boss," no cycles, and concepts that are held together primarily by subordinate relations.

Defining the structure of a particular content domain, then, is not an easy task. Different conclusions about the structure may arise if different experts and criteria are used.

In this study, we examined the intersection of task demands and constraints with the structure of the subject domain to be mapped. Two mapping techniques with the same task demand (i.e., construct a map), but different task constraints (i.e., imposing on students a specific structure for their maps) were used. Mapping Technique 1 asked students to construct their concept maps using a hierarchical structure, and Mapping Technique 2 asked students to construct their maps organizing the concepts in any way they wanted. To evaluate the intersection of imposing a structure with the structure of the subject domain, two content domains were selected, one with a "hierarchical" structure and another one with a "non-hierarchical" structure as determined by two experts. If the structure of a map reflects the structure in the domain as held by an expert, we expected that students who knew the subject matter would construct maps with similar structures to that of the expert's.

Method

Participants

Two classes of high school chemistry students taught by the same teacher (with 7 years of teaching experience), a second chemistry teacher (with 5 years of teaching experience), and two experts, one a chemist (with 10 years of experience in research on water quality) and the other a physicist (with 14 years of experience in research on subatomic particles) participated in the study. The students, the teachers and the experts were trained to construct concept maps with the same training program. All subjects were drawn from the Palo Alto area.

Of the original 62 students in the two groups, 8 students were dropped from the data set because of incomplete data. Another 6 students were randomly dropped to provide pilot data to check out scoring procedures, producing equal cell sizes. As a result, data for 48 students were analyzed.

Design

Two topics were selected as having different structures according to the experts' maps: one topic with a hierarchical content structure and another one with a non-hierarchical content structure. Classes were randomly assigned to the topic in which they were assessed. Within each class students were randomly assigned to one of two mapping techniques: Mapping Technique 1—instructions imposing the construction of hierarchical maps (Hierarchical Instructions) and Mapping Technique 2—instructions without restrictions on the type of structure for constructing their maps (Non-Hierarchical Instructions). This factorial design had two between-subjects factors: (1) Topic, with two levels: topic with hierarchical structure and topic with non-hierarchical structure; and (2) Mapping technique, with two levels: Hierarchical Instructions and Non-Hierarchical Instructions.

Domain and Material

The two topics selected for this study were "Atomic Structure" and "Nature of Ions, Molecules, and Compounds" which are topics involved in the big idea, "Reactions and Interactions," as described in the Science Framework for California Public Schools (California Department of Education, 1990). These two topics were taught as two consecutive units in the chemistry curriculum at the school where the study was conducted. Both units were taught using the chapters "Atom Structure" and "Chemical Names and Formulas" from the widely used textbook, *Chemistry* (Wilbraham, Staley, Simpson, & Matta, 1990).

The two experts were used to define the content structure of the two topics. According to their area of expertise, the two experts were asked to construct a concept map on either “Atom Structure” or “Ions, Molecules, and Compounds.” The “hierarchiness” of the experts’ maps were judge based on four aspects: (a) the number of “beginner” nodes (i.e., nodes with only arrows coming-out but no arrows coming-in), (b) the number of nodes with more than “one boss” (i.e., nodes with more than one arrow coming into the node), (c) the number of cycles or “loops” in the map, and (d) the percentage of subordinate propositions. The expert’s map on the topic “Atom Structure” had one “beginner” node; three nodes with more than one arrow coming in; no cycles; and 95 percent of the propositions in this map were subordinate. The expert’s map for “Ions, Molecules, and Compounds” had three “beginner” nodes; ten nodes with more than one arrow coming-in; no cycles; and 97 percent of the propositions in his map were subordinate. Based on this information, the topic, “Atom Structure,” was considered as having a more hierarchical structure than the topic, “Ions, Molecules, and Compounds.”

The chapters “Atom Structure” and “Chemical Names and Formulas,” were used to define the content domain for selecting the concepts used in the study. We compiled two list of key concepts by: (1) asking the chemistry teachers to provide the concepts they thought were most important in the unit; (2) asking the experts to provide the concepts they thought were the most important for students to know, based on the content provided in the chapters; and (3) reviewing the textbook used in the class ourselves. The procedure for sampling concepts is described in Appendix A. A list of 17 key concepts was compiled from the “Atom Structure” chapter. For the chapter, “Chemical Names and Formulas,” we used the list compiled for a previous study (i.e., Ruiz-Primo, Schultz, & Shavelson, 1996; Ruiz-Primo & Shavelson, 1997) and eliminated, from the 20 key concept list, the three concepts (i.e., binary molecular compounds, negative charge, and positive charge) that had the least number of connections with other concepts based on a criterion map. We had, then, two 17-key concept lists, one for each topic used in the study (see Appendix B).

Instrumentation

Concept map task. The two mapping techniques explored in this study varied in the task constraints imposed on the students: constructing a hierarchical or non-hierarchical map. Mapping technique 1—hierarchical structure imposed—asked students to construct a 17-concept map organizing the more general terms above more specific terms (see Appendix C). Mapping technique 2—no specific structure imposed—asked students to construct a 17-concept map organizing the terms in any way they wanted (see Appendix C).

Scoring system. The scoring system was based on a *criterion map*—a composite of the experts', teachers', and researchers' maps. Two 17-concept criterion maps were constructed to identify those propositions "substantial" to the domain, and that students should know about at that point in the chemistry course.

Based on the 17 key concepts, a square-matrix was constructed to define all possible links between concept pairs. The entry in a cell of the matrix denoted the relation between a specific concept pair. Up to 136 links could be drawn between the pairs of the 17 concepts. To determine the "substantial" links, teachers, experts, and researchers constructed concept maps. Teachers' concept maps were expected to provide a benchmark for the "substantial" links students were expected to have after studying the chapters and participating in class. The experts' concept maps provided the "substantial" links based on the structure of the discipline. Finally, researchers' concept maps were thought to reflect the "substantial" links in the textbook chapters. Propositions were carefully analyzed across the three maps for each topic to determine whether a particular relation should be included in the criterion map. The propositions that were the same across the maps were considered "mandatory"—students should be reasonably expected to provide any one of these propositions at that point in their instruction. The analysis of the maps constructed by the teachers, the experts, and the researchers identified 25 "mandatory" propositions for the topic, "Atom Structure," and 44 "mandatory" propositions for the topic, "Ions, Molecules and Compounds."

To account for the variation in the quality of the propositions, we developed for each topic a *Proposition Inventory*. Each inventory compiled the propositions (nodes and direction of links) provided by the teachers' maps, experts' maps, students' maps and researchers' maps and classified each proposition into one of five categories: Accurate Excellent, Accurate Good, Accurate Poor, Don't Care, and Inaccurate. Table 3 presents the definition of each category. For example, the accurate excellent proposition between *acids* and *compounds* should be read, according to the direction of the arrow (<), as follows: *compounds* that give off H+

when dissolved in water are *acids*. Both inventories were judged by the experts to determine the accuracy/validity of the classification of the propositions. Few changes were necessary.

The Proposition Inventories provided propositions not only considered "mandatory," but also propositions for the "other" possible relations between the pairs of concepts in the key concept list. These other propositions were considered as *"possible"* propositions. Both inventories included the "mandatory" and the "possible" propositions (i.e., the 136 possible links between the pairs of the 17 key/core concepts). Therefore, students were credited for any valid proposition that they provided that was not contained on the criterion map.

Table 3
Accuracy of the Propositions

Accuracy of Proposition	Definition
Excellent:	Outstanding proposition. Complete and correct. It shows a deep understanding of the relation between the two concepts. *acids-compounds: < that gives off* H^+ *when dissolved in water are*
Good:	Complete and correct proposition. It shows a good understanding of the relation between the two concepts. *acids-compounds: > are examples of*
Poor:	Incomplete but correct proposition. It shows partial understanding of the relation between the two concepts. *acids-compounds: < form*
Don't Care:	Although valid, the proposition does not show understanding between the two concepts. *acids-compounds: > is a different concept*
Inaccurate/ Invalid:	Incorrect proposition. *acids-compound: > made of*

The scoring system, based on the criterion map and the Propositions Inventory, evaluated two aspects of the map: the propositions and the structure.

The accuracy of each proposition in a student's map was assessed on a five-level scale (from 0 for inaccurate to 4 for accurate excellent) according to the Propositions Inventory. Three concept map scores were formed: (1) a total *proposition accuracy* score—the total sum of the scores obtained on all propositions; (2) *convergence* score—the proportion of valid propositions in the student's map out of all mandatory propositions in the criterion map (i.e., the degree to which the student's map and the criterion map converge); (3) *salience* score—the proportion of valid propositions out of all the propositions in the student's map.

To score the "hierarchiness" of the map structures, we evaluated four aspects of the student's map structure as to whether: (1) the map had only *one* "beginner or commander" node (i.e., a node that had only arrows coming out, but *none* coming in), (2) the nodes had a "single boss" (i.e., on each node only *one* arrow comes in), (3) the relations among nodes were asymmetrical and transitive (i.e., no cycles in the structure), and (4) the relations between node pairs were subordinate (i.e., for each proposition, one of the nodes is considered to be less general, or have an inferior rank, or be under the control of the other). Information provided by these four aspects bear, directly or indirectly, on the five criteria proposed by Bunge for classifying a structure as hierarchical.

Separate score forms were designed for each topic. Appendix D shows the form used to score the Atom Structure concept maps. Two raters scored each student map.

Multiple-Choice Test. Prior to administering the concept maps, both classes received a 15-item multiple-choice test: Group 1 received the test on "Atom Structure" and Group 2 on "Ions, Molecules, and Compounds." The multiple-choice tests were designed by the researchers and reviewed by the teachers. The internal consistency reliability of the "Atom Structure" test was .56, and .71 for the "Ions, Molecules, and Compounds" test. Three unrelated items were dropped from the atom structure test to increase the internal consistency coefficient to .56.

Training

A program was designed to teach students, teachers, and experts to construct concept maps. The program was piloted and evaluated with another two groups of high school chemistry students. Based on the pilot minor modifications were made. The same researcher trained both groups of study students to minimize variability.

The training lasted about 50 minutes and had four major parts. The first part focused on introducing concept maps: what they are, what they are used for, what their components are (i.e., nodes, links, linking words, propositions), and examples (outside the domain to be mapped) of hierarchical and non-hierarchical maps. The second part emphasized the construction of concept maps. Four aspects of mapping were highlighted: identifying a relationship between a pair of concepts, creating a proposition, recognizing good maps, and redrawing a map. Students were then given two lists of common concepts to “collectively construct” a map. The first list focused on the “water cycle”—a non-hierarchical map; the second list focused on “living things”—a hierarchical map. The third part of the program provided each individual with 9 concepts on the “food web,” and asked them to construct a map individually. The fourth part of the program was a discussion of students’ questions after they had constructed their individual maps.

To evaluate the training a random sample of 10 of the individually constructed maps was analyzed for each group (a total of 20 concept maps). Results indicated that: (a) 100 percent of the students in Group 1 (those who studied the Atom Structure topic) and 97.8 percent in Group 2 (those who studied the Ions, Molecules, and Compounds topic) used all the concepts provided on the list, (b) 100 percent of the students in both groups used labeled lines, and (c) 85.6 and 89.5 percent of the students’ propositions, in Group 1 and 2 respectively, were valid. We concluded that the training program succeeded in training the students to construct concept maps.

Procedure

The study was conducted in three 55-minute sessions during a four-week period. The first session was used for training, the second for the multiple-choice test. In the third session, students received, first, a 15-minute reminder on how to construct concept maps and then they were asked to construct the concept map.

Both classes were trained on the same day in their respective classrooms before the unit, “Atom Structure,” was taught. Group 1 had the second and third sessions two weeks after the training, when the instruction on the unit, “Atom Structure,” ended. Group 2 received the sessions four weeks after the training, when the instruction of the unit, “Chemical Names and Formulas,” ended. Construction of concept maps took 25 minutes, on average, for both groups.

Results

This study addressed the questions: Is there an effect of imposing a hierarchical structure (Mapping Technique 1) and non-hierarchical structure (Mapping Technique 2) on students’ representations of two types of content domains? Do concept maps provide reliable scores? Do concept maps provide information about student knowledge similar to multiple-choice tests?

Following the 2x2 factorial design the characteristics of the four groups were: Group 1—students mapped the hierarchically structured topic (i.e., Atom Structure) and received Mapping Technique 1 with instructions imposing a hierarchical structure for constructing their maps (HT/HI). Group 2—students mapped the hierarchically structured topic and received Mapping Technique 2 with instructions not restricting the type of structure for constructing their maps (HT/NHI). Group 3—students mapped the non-hierarchically structured topic (i.e., Ions, Molecules, and Compounds) and received Mapping Technique 1 (NHT/HI). And Group 4—students mapped the non-hierarchically structured topic and received Mapping Technique 2 (NHT/NHI).

Groups were compared using the three types of scores based on the quality of the propositions and the hierarchiness of the students’ map structures: Is there an instruction effect in students’ map scores? Do students’ map structures differ in the degree of hierarchiness depending on the mapping technique and the structure of the topic? (e.g., Was the hierarchically structured topic mapped in a hierarchical way even though instructions to do so were not provided?).

Concept Map Scores

Comparing Means. Table 4 shows the means and standard deviations for proposition accuracy, convergence and salience scores across groups. Mean of the proposition accuracy scores across the groups revealed that student’s knowledge was partial and not close to the standard established by the criterion maps. Lower mean scores observed for groups 3 and 4 indicated that students' knowledge about "Ions, Molecules, and Compounds" was weaker than students' knowledge about "Atom Structure" when compared with the criterion maps. The same pattern is observed in the convergence and salience mean scores.

It is important to note that the topic, "Ions, Molecules, and Compounds" was more complex than the topic, "Atom Structure." The number of mandatory links in the Ions, Molecules, and Compounds criterion map almost doubled that of Atoms Structure.

Table 4
Mean and Standard Deviations of the Propositions Accuracy, Convergence, and Salience on Each Condition

Group	Proposition Accuracy		Convergence		Salience	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Atoms						
1 HT/HI	36.92 (Max=100)	13.66	0.51 (Max P=25)	0.16	0.76	0.15
2 HT/NHI	40.29 (Max=100)	9.64	0.54 (Max P=25)	0.11	0.78	0.15
Ions, Molecules, & Compounds						
3 NHT/HI	28.29 (Max =176)	15.73	0.23 (Max P=44)	0.12	0.57	0.27
4 NHT/HI	28.99 (Max =176)	15.14	0.23 (Max P=44)	0.11	0.60	0.25

[a] Maximum score was calculated based on 25 excellent valid mandatory propositions students could provide.

[b] Maximum score was calculated based on 44 excellent valid mandatory propositions students could provide.

To evaluate the interaction of topic and mapping technique three 2x2 factorial ANOVAs were carried out, one for each type of score. To adjust the number of expected propositions in both topics, we transformed the proposition accuracy score into a proportion (i.e., student's total score divided by the total of all mandatory propositions).[5] Results for the proposition accuracy score indicated no significant interaction of topic by mapping technique or mapping technique main effect (F_{TxMT} = .25; p > .05; F_{MT} = .39; p > .05), and, not surprisingly, a significant topic main effect (F_T = 55.26; p < .05), although this result is not of special interest for our purposes. ANOVA results for convergence and salience scores also found no significant interaction or mapping technique effect and a significant topic effect (Convergence: F_{TxMT} = .072, F_{MT} = .103, p > .05; and F_T = 63.61; p < .05; and Salience: F_{TxMT} = .004, F_{MT} = .193, p > .05; and F_T = 9.13; p < .05).

Comparing Types of Scores. To examine the convergence of the scores, we created a multiscore matrix for each topic, Atom Structure and Ions Molecules & Compounds (Table 5). Interrater reliability coefficients are enclosed by parenthesis on the diagonal. Coefficients across types of score indicate that raters can consistently score concepts maps.

Table 5

Multiscore Matrix

	Atom Structure			Ions, Molecules and Compounds		
	PA	C	S	PA	C	S
Proposition Accuracy (PA)	(.98)			(.99)		
Convergence (C)	.95	(.98)		.95	(.99)	
Salience (S)	.86	.88	(.95)	.91	.95	(.99)

Coefficients are high and roughly the same across topics.[6] This suggest that proposition accuracy, convergence and salience scores, in general, rank students

---

[5] Although raw scores are presented in Table 4, all statistical analyses for proposition accuracy, convergence and salience scores were carried out using both the raw scores and the arcsin transformation of the proportions. Both analyses provided the same results across the three types of scores.

[6] A multiscore matrix for each group shows same patterns and similar magnitudes of the interrater and convergence coefficients presented in Table 5, except for Group 2 (HT/NHI). The Multiscore matrix for Group 2 shows a convergence coefficient between proposition accuracy and salience score (.90) a little bit higher than the coefficient for the convergence and salience scores (.89).

similarly. However, the highest correlations were the correlations between proposition accuracy and convergence scores ($\underline{r}_{avg.}$ = .95), followed by the convergence and salience score correlations ($\underline{r}_{avg.}$ = .92), and then by the propositions accuracy and salience score correlations ($\underline{r}_{avg.}$ = .86). This suggests that the type of method selected for scoring concept maps might be an issue since not all of them rank students exactly in the same way. Proposition accuracy and convergence scores seems to rank more consistently than salience scores.

Relation to Multiple-Choice Tests. To evaluate the extent to which concept maps measure different aspects of declarative knowledge than multiple-choice tests, the correlation between these two measures was calculated. Correlation coefficients are presented in Table 6. Along with the original correlations, we present correlations corrected for attenuation using the interrater reliability coefficient for each score. (This correction may not be accurate and must be interpreted cautiously. Reliability coefficients involved in the correction are not equivalent because the type of errors of measurement differ for each reliability coefficient.)

Table 6

Correlation Between the Multiple-Choice Test and Proposition Accuracy, Convergence, and Salience Score by Topic

Topic	Proposition Accuracy	Convergence	Salience
Atom Structure			
Observed	.39	.27	.18
Disattenuated	.53	.36	.34
Ions, Molecules, & Compounds			
Observed	.36	.32	.35
Disattenuated	.43	.38	.42

We focused on the original coefficients which are positive and not very high. We interpret these coefficients to mean that both tests measured the same knowledge domain, but still different aspects of it. However, coefficients across types of scores and topics are not the same. For the Atom Structure, coefficients are higher for proposition accuracy and convergence scores than salience scores. However, for Ions, Molecules and Compounds, coefficients are more consistent across types of scores.

Concept Map Hierarchiness

The "hierarchiness" of the students' maps was evaluated using the following data: the number of beginner-nodes, the number of cycles, the number of circuits, and the proportion of subordinate propositions out of all the propositions in the student's map.

Only the maps of the top 25 percent of the students on each group based on the concept map scores were evaluated for "hierarchiness." Those students with low scores did not have sufficient knowledge to reflect the structure of the content domain. The mean scores of the four students from each group are presented in Table 7.

Table 7

Mean and Standard Deviations of the Propositions Accuracy, Convergence, and Salience Scores Considering Only the Four Top Students on Each Group.

Group	Proposition Accuracy		Convergence		Salience	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Atoms						
1 HT/HI	51.25 (Max=100)	7.60	0.66 (Max P=25)	0.09	0.87	0.02
2 HT/NHI	49.75 (Max=100)	5.19	0.63 (Max P=25)	0.07	0.94	0.07
Ions, Molecules, & Compounds						
3 NHT/HI	44.38 (Max =176)	9.93	0.35 (Max P=44)	0.05	0.80	0.04
4 NHT/HI	44.88 (Max =176)	10.04	0.32 (Max P=44)	0.09	0.77	0.11

[a] Maximum score was calculated based on 25 excellent valid mandatory propositions students could provide.

[b] Maximum score was calculated based on 44 excellent valid mandatory propositions students could provide.

The pattern of mean scores observed for all students is similar to this selected group of students; means were higher for Groups 1 and 2 on the three types of scores. However, students' maps still indicated partial knowledge about the topics when compared with the criterion maps.

The "hierarchiness" of the top-students' maps were scored by two raters. Interrater reliability coefficients across hierarchiness scores were also high: .82 for beginner nodes; 1.00 for cycles; .98 for circuits, and .87 for subordination. Table 8 presents the mean for each hierarchiness score.

To evaluate whether an interaction effect—topic by mapping technique—was observed, a factorial ANOVA was carried out for each of the hierarchiness scores. No significant interaction or main effect ($p > .05$) was found in any of the "hierarchiness" scores (Beginner Nodes: $F_{TxMT} = .67$, Circuits: $F_{TxMT} = .009$, and Subordination: $F_{TxMT} = .34$). These results indicated that imposing a hierarchical structure does not interact with the structure of the content domain mapped. However, this interpretation seems premature since some problems arose in the way "hierarchical structure" was defined. For example, according to the four

Table 8

Mean and Standard Deviations of the "Hierarchiness" Scores on Each Group Considering Only the Four Top Students

	Beginner Nodes		Cycles		Circuits		Subordinate	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Atom								
1 HT/HI	3.13	3.13	0	0	2.50	1.29	0.82	0.08
2 HT/NHI	3.62	2.28	0	0	1.75	0.50	0.83	0.04
Ions, Molecules, & Compounds								
3 NHT/HI	4.13	1.18	0	0	2.75	1.70	0.79	0.13
4 NHT/NHI	3.25	1.89	0	0	1.87	1.55	0.89	0.11

"hierarchiness" criteria used, no pure "hierarchical" structures could be identified in any of the students' maps. However, the high proportion of subordinate relations suggest a high degree of hierarchiness in all the students' maps, independent of the condition. Furthermore, when only subordinate scores are considered, a completely different picture would emerge: all maps could be considered as hierarchical since most of the relations between concepts are held in superordinate/subordinate relation.

A closer examination of the criteria used to define "hierarchiness" and further analysis of the students' maps (e.g., analysis of the characteristics of the levels in the student's maps) are needed before any final decision is made about the use of hierarchical instructions for constructing concept maps.

Conclusion

This study explored two concept mapping techniques that varied instructions provided to students for constructing their maps. We examined: (1) whether imposing a structure on students' representations interact with the structure of the subject domain mapped; (2) how consistent were map score across raters; and (3) whether concept map-based assessments provide information about students' knowledge similar to multiple-choice tests.

Final conclusions are not possible at this time because of the issues involved in defining the structure of the subject domain to be mapped and the criteria used to define hierarchiness. Nevertheless, some things were learned: (1) Findings about interrater reliability are encouraging. Concept maps can be reliably scored despite the complex judgments involved in assessing the quality of students' propositions. Interrater reliability coefficients were typically high (above .90) across the three types of scores. (2) Correlations between multiple-choice tests and concept map scores across types of scores were all positive and moderate ($\underline{r}$ = .31 on average) suggesting that both types of assessment measure overlapping and yet somewhat different aspects of declarative knowledge. (3) The type of score selected for scoring concept maps might be an issue. Proposition accuracy and convergence scores seems to be more consistent than the salience score. These results confirm findings in a previous study (e.g., Ruiz-Primo & Shavelson, 1997): salience scores ranked students differently according to the mapping technique used. Furthermore, based on the results of a person x rater x concept sample G study, it was found that the percent of variability among persons was higher for the proposition accuracy and convergence score than for the salience score. This indicated that these two types of scores reflect better the differences in students' knowledge structure than salience scores.

Unfortunately, criteria used to define hierarchiness prevent us from arriving at a final conclusion about the interaction between hierarchical instructions and the structure of the subject-matter domain. We recognize that different conclusions about the structure of the topics and the students' map structures could arise if different experts and "hierarchiness" criteria were used. It may be that an "averaged" experts' structure should be consider for defining the structure of the domain. This may reduce the problem of variability among experts and provide a better picture of the structure of the content domain (e.g., Acton, Johnson and Goldsmith, 1994). Which criteria (e.g., subordination, characteristics of levels in the structure, hierarchical clusters) should be used to define a hierarchy is a compelling research question worth further exploration. For now, we are exploring cluster analysis for further evaluation of students' maps.

Many questions still remain to be studied. The research agenda in this area is extensive (see Ruiz-Primo & Shavelson, 1996) but necessary if we want to test the potential of concept maps as an alternative assessment in science.

References

Acton, W. H., Johnson, P. J., Goldsmith, T. E. (1994). Structural knowledge assessment: Comparison of referent structures. *Journal of Educational Psychology, 86*(2), 303-311.

Anderson, R. C. (1984). Some reflections on the acquisition of knowledge. *Educational Researcher, 13*(10), 5-10.

Anderson, T. H. & Huang, S-C. C. (1989). *On using concept maps to assess the comprehension effects of reading expository text* (Technical Report No. 483). Urban-Champaign: Center for the Studying of Reading, University of Illinois at Urban-Champaign. (ERIC Document Reproduction Service No. ED 310 368).

Barenholz, H. & Tamir, P. (1992). A comprehensive use of concept mapping in design instruction and assessment. *Research in Science & Technological Education, 10*(1), 37-52.

Bunge, M. (1969). The metaphysics, epistemology and methodology of levels. In L. L. Whyte, A. G. Wilson, & D. Wilson (Eds.), *Hierarchical structures* (pp. 17-28). New York: American Elsevier Publishing Company, Inc.

California Department of Education (1990). *Science framework for California Public Schools: Kindergarten through grade twelve.* Sacramento, CA: California Department of Education.

Cronbach, L. J. (1990). *Essentials of psychological testing* (Fifth ed.). New York: Harper & Row Publishers.

Dochy, F. J. R. C. (1996). Assessment of domain-specific and domain-transcending prior knowledge: Entry assessment and the use of profile analysis. In M. Birenbaum & F. J. R. C. Dochy (Eds.), *Alternatives in assessment of achievements, learning process and prior knowledge* (pp. 93-129). Boston, MA: Kluwer Academic Publishers.

Dress, A. & von Haeseler, A. (1990). Trees and hierarchical structures. In S. Levin, *Lecture Notes in Biomathematics*, 84. Berlin: Springer-Verlag.

Fisher, K. M. (1990). Semantic networking: The new kid on the block. *Journal of Research on Science Teaching, 27*, 1001-1018.

Glaser, R. (in press). Changing the agency for learning: Acquiring expert performance. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of* *expert performance in the art, and sciences, sports, and games*. Mahwah, NJ: Erlbaum.

Glaser, R. & Bassok, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology, 40*, 631-66.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology, 83*(1), 88-96.

Green, M. (1969). Hierarchy: One word, How many concepts? In L. L. Whyte, A. G. Wilson, & D. Wilson (Eds.), *Hierarchical structures* (pp. 56-58). New York: American Elsevier Publishing Company, Inc.

Harnisch, D. L., Sato, T., Zheng, P., Yamaji, S. & Connell, M. (in press). Concept mapping approach and its implications in instruction and assessment. *Computers in School*.

Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge. Techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Jones, R. G. (1969). Structural hierarchy: The problem of the one and the many. In L. L. Whyte, A. G. Wilson, & D. Wilson (Eds.) *Hierarchical structures* (pp. 56-58). New York: American Elsevier Publishing Company, Inc.

Lomask, M., Baron, J. B., Greig, J. & Harrison, C. (1992, March). *ConnMap: Connecticut's use of concept mapping to assess the structure of students' knowledge of science*. Paper presented at the annual meeting of the National Association of Research in Science Teaching. Cambridge, MA.

Markham, K. M., Mintzes, J. J., & Jones, M. G. (1994). The concept map as a research and evaluation tool: Further evidence of validity. *Journal of Research in Science Teaching, 31*(1), 91-101.

Mesarovic, M. D., & Macko, D. (1969). Foundations for a scientific theory of hierarchy systems. In L. L. Whyte, A. G. Wilson, & D. Wilson (Eds.) *Hierarchical structures* (pp. 29-50). New York: American Elsevier Publishing Company, Inc.

Nakhleh, M. B. & Krajcik, J. S. (1991). *The effect of level of information as presented by different technology on students' understanding of acid, base, and pH concepts*. Paper presented at the annual meeting of the National Association for the Research in Science Teaching, Lake Geneva, WI. (ERIC Document Reproduction Service No. ED 347 062).

Novak, J. D., Gowin, D. B., & Johansen, G. T. (1983). The use of concept mapping and knowledge vee mapping with junior high school science students. *Science Education, 67*(5), 625-645.

Rosen, R. (1969). Comments on the use of the term hierarchy. In L. L. Whyte, A. G. Wilson, & D. Wilson (Eds.) *Hierarchical structures* (pp. 56-58). New York: American Elsevier Publishing Company, Inc.

Ruiz-Primo, M. A. & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching, 33*(6), 569-600.

Ruiz-Primo, M. A. & Shavelson, R. J. (1997) *Concept-map based assessment: On possible source of sampling variability*. Manuscript submitted for publication.

Ruiz-Primo, M. A., Schultz, S. E. & Shavelson, R. J. (1996, April). Concept map-based assessment: An exploratory study. Paper presented at the AERA Annual Meeting, New York, NY.

Roth, W. M. & Roychoudhury, A. (1993). The concept map as a tool for the collaborative construction of knowledge: A microanalysis of high school physics students. *Journal of Research in Science Teaching, 30*(5), 503-534.

Shavelson R. J. (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology, 63*, 225-234.

Wallace, J. D. & Mintzes, J. J. (1990). The concept map as a research tool: Exploring conceptual change in biology. *Journal of Research in Science Teaching, 27*(10), 1033-1052.

White, R. T, & Gunstone, R. (1992). *Probing understanding*. New York: Falmer Press.

Whyte, L. L., Wilson, A. G., & Wilson, D. (Eds.). (1969). *Hierarchical structures*. New York: American Elsevier Publishing Company, Inc.

Wilbraham, A. C., Staley, D. D., Simpson, C. J., & Matta, M. S. (1990). *Chemistry*. Menlo-Park, CA: Addison-Wesley Publishing Company.

Appendix A

Compiling the List of Concepts

Teachers were asked to answer these two questions about each unit: (1) "Explain in a few words what you want your students to know when they finish this chapter. In answering this question, think about why this chapter is included in the curriculum and what is the most important thing you want the students to learn about the topic"; and (2) "Based on your answer to question 1, please review the chapter and list all the concepts you think students should know and understand after studying this topic."

Teachers' answers to question 1 involved two aspects of the unit: (a) *conceptual understanding* (e.g., "Students should have a good understanding of the formation of ions, the differences between molecules/compounds... molecular/ionic compounds and acids.") and (b) *application* (e.g., "They should be able to form ionic compounds, binary, ternary, and acids...be familiar with the periodic table to identify metals, non-metals...." "Students... should be able to write chemical/molecular formulas; name different substances..."). We focused on the conceptual understanding of the unit since concepts maps are about the interrelatedness of concepts.

Our list included 14 and 23 key concepts selected from the "Atom Structure and Chemical Names and Formulas" chapters, respectively. We gave this list to the teachers with the following instructions: "This is a list of concepts that were selected from the chapter, "X." Based on what you think are the most important ideas for students to understand about "X," check (√) the concepts that are essential. Please feel free to add any concepts that are missing." Based on the concepts selected and added by the teacher we increased and reduced the lists to 17 concepts each (see Appendix B). These two lists of concepts were considered to represent the "key concepts" of the chapters.

Appendix B

Concept Lists: Key Concept List, Researchers' List, Teachers' List, and Expert's List

Atom Structure

Key Concept List	Researcher's List	Teachers' List	Physicist's List
1. atom	1. atom	1. atom	1. atom
2. atomic mass	2. atomic mass	2. atomic mass	2. atomic mass
3. atomic number	3. atomic number	3. atomic number	3. atomic number
4. atomic orbitals	4. electron	4. *d* orbital	4. binding energy
5. electron	5. isotope	5. Dalton's atomic theory	5. electromagnetic force
6. elements	6. mass number	6. electron	6. electron
7. energy levels	7. negative charge	7. element	7. filled orbitals
8. isotope	8. neutral charge	8. energy levels	8. ions
9. mass number	9. neutron	9. isotope	9. isotope
10. negative charge	10. nucleus	10. mass number	10. mass number
11. neutral charge	11. orbitals	11. negative charge	11. neutron
12. neutron	12. positive charge	12. neutral charge	12. nucleus
13. nucleus	13. proton	13. neutron	13. electron cloud (orbitals)
14. *p* orbitals	14. subatomic particles	14. nucleus	14. Pauli exclusion principle
15. positive charge		15. orbitals	15. periodic table
16. proton		16. *p* orbitals	16. proton
17. *s* orbitals		17. periodic table	17. photoelectric effect
		18. positive charge	18. quarks
		19. proton	19. shape of orbitals
		20. *s* orbitals	20. strong force
			21. subatomic particles
			22. unfilled orbitals
			23. weak force

Ions, Molecules, and Compounds

Key Concept List	Researcher's List	Teachers' List	Physicist's List
1. acids	1. acids	1. acids	1. acids
2. anions	2. anions	2. anions	2. anions
3. atoms	3. atoms	3. cations	3. atoms
4. bases	4. bases	4. compounds	4. bases
5. binary ionic compounds	5. binary ionic compounds	5. element	5. binary ionic compounds
6. cations	6. binary molecular compounds	6. ionic charge	6. cations
7. compounds	7. cations	7. ionic compounds	7. compounds
8. electrons	8. compounds	8. molecules	8. electrons
9. ions	9. electrons	9. molecular compounds	9. elements
10. ionic compounds	10. elements	10. periodic table	10. ions
11. metals	11. ions	11. chemical formulas	11. ionic compounds
12. molecules	12. ionic compounds	12. molecular formulas	12. metals
13. molecular compounds	13. metals		13. molecules
14. neutral charge	14. metalloids		14. molecular compounds
15. non-metals	15. molecules		15. negative charge
16. polyatomic ions	16. molecular compounds		16. neutral charge
17. ternary ionic compound	17. negative charge		17. non-metals
	18. neutral charge		18. representative elements
	19. non-metals		19. polyatomic ions
	20. polyatomic ions		19. positive charge
	21. positive charge		20. ternary ionic compound
	22. ternary ionic compound		21. transition elements
	23. transition metals		

Appendix C

Sample of Hierarchical and Non-Hierarchical Instructions

Mapping Technique 1 Instructions—Hierarchical Structure is Imposed.

Name ____ Period________

Examine the concepts listed below. They were selected from the chapter on Atomic Structure that you recently studied. Construct a hierarchical concept map using the terms provided below. Organize more general terms above the more specific ones. Draw a line between the terms you think are related. Label the line using phrases or only one or two words.

You can construct your map on the blank pages attached. When you finish your map check that: (1) you have all the concepts on the list in your map; (2) all the lines have labels; (3) your map is explaining atomic structure. After checking your map <u>redraw</u> it so someone else can read it.

Staple your <u>final map</u> to this page.

LIST OF CONCEPTS
atoms
atomic mass
atomic number
atomic orbitals
electrons
elements
energy levels
isotopes
mass number
negative charge
neutral charge
neutrons
nucleus
p orbitals
positive charge
protons
s orbitals

Concept Mapping Technique 2 Instructions—No Specific Structure is Imposed.

Name ____ Period________

Examine the concepts listed below. They were selected from the chapter on Atomic Structure that you recently studied. Construct a concept map using the terms provided below. Organize the terms in relation to one another in any way you want. Draw a line between the terms you think are related. Label the line using phrases or only one or two words.

You can construct your map on the blank pages attached. When you finish your map check that: (1) you have all the concepts on the list in your map; (2) all the lines have labels; (3) your map is explaining atomic structure. After checking your map <u>redraw</u> it so someone else can read it.

Staple your <u>final map</u> to this page.

LIST OF CONCEPTS
atoms
atomic mass
atomic number
atomic orbitals
electrons
elements
energy levels
isotopes
mass number
negative charge
neutral charge
neutrons
nucleus
p orbitals
positive charge
protons
s orbitals

Appendix D

Atom Structure Scoring Form

Name_____________________________________ ID____________________

Date ______________________ Rater ______________________

Mandatory		
# R	VS	IS
2		
3		
5		
7		
11		
22		
33		
35		
40		
42		
45		
49		
57		
60		
66		
72		
92		
93		
99		
101		
107		
109		
110		
125		
134		

Possible		
# R	VS	IS
1		
4		
6		
8		
9		
10		
13		
15		
16		
18		
19		
20		
23		
24		
30		
32		
34		
46		
47		
48		
50		
55		
58		
59		
65		
67		
69		
70		
71		
73		
74		
75		
76		
78		
80		
81		
88		
91		
95		
103		
123		
126		
127		
130		
133		

No-Relation		
# R	VS	IS
17		
21		
25		
26		
27		
28		
29		
31		
36		
37		
38		
39		
41		
43		
44		
51		
52		
53		
54		
56		
61		
62		
63		
82		
83		
84		
85		
86		
87		
89		
90		
94		
96		
97		
98		
100		
102		
104		
105		
106		
108		
112		
114		
115		
118		
121		
128		
132		
136		

Forbidden		
# R	VS	IS
12		
14		
64		
68		
77		
79		
111		
113		
116		
117		
119		
120		
122		
124		
129		
131		
135		

Comments:

____

# of Concepts: ______

atoms ______
atomic orbitals ______
atomic mass ______
atomic number ______
electrons ______
elements ______
energy level ______
isotopes ______
mass number ______
neutrons ______
neutral charge ______
nucleus ______
negative charge ______
p orbitals ______
positive charge ______
protons ______
s orbitals ______

Valid Prop ______

Total # of Prop ______

Inclusive Rel. ______

of Cycles ______

# of Circuits ______

# of Beg Nodes ______

Total M [ ] Total P [ ] Total NR [ ] Total F [ ]

Grand Total [ ]