Third Misconceptions Seminar Proceedings (1993)

Paper Title: Assessing Conceptual Understanding in Science through the Use of Two- and Three Dimensional Concept Maps

Author: Lomask, Michal S.; Baron, Joan B. & Grieg, Jeffrey

Abstract: This paper describes a method to analyze students' understanding of scientific concepts, by converting essay-type responses into two- and three-dimensional concept maps. In this method, a student's written response is converted into a concept map and then compared to concept maps created by experienced science teachers. The relevancy and validity of the concepts are the main attributes of the rating, which is used to report the size and strength of students' structures of knowledge. The paper describes the concept mapping scoring method, shows examples of students' work (including three-dimensional group concept maps) and discusses various aspects of this method.

Keywords:
General School Subject:
Specific School Subject:
Students:

Macintosh File Name: Lomask - 3D Concept Maps
Release Date: 12-16-1993 C, 11-6-1994 I

Publisher: Misconceptions Trust
Publisher Location: Ithaca, NY
Volume Name: The Proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics
Publication Year: 1993
Conference Date: August 1-4, 1993
Contact Information (correct as of 12-23-2010):
Web: www.mlrg.org
Email: info@mlrg.org

A Correct Reference Format: Author, Paper Title in The Proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, Misconceptions Trust: Ithaca, NY (1993).

you have found this collection to be of value in your work, consider supporting our ability to support you by purchasing a subscription to the collection or joining the Meaningful Learning Research Group.

----- -----

Assessing Conceptual Understanding in Science through the Use of Two- and Three Dimensional Concept Maps

By
Michal S. Lomask
Joan B. Baron
Jeffrey Grieg

Connecticut State Department of Education

Paper presented at the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics

Cornell University, Ithaca, New York. August 1-4, 1993

Assessing Conceptual Understanding in Science through the Use of Two- and Three-Dimensional Concept Maps

Michal S. Lomask and Joan Boykoff Baron, Jeffrey Grieg,

Connecticut State Department of Education

ABSTRACT

This paper describes a method to analyze students' understanding of scientific concepts, by converting essay-type responses into two- and three-dimensional concept maps. In this method, a student's written response is converted into a concept map and then compared to concept maps created by experienced science teachers. The relevancy and validity of the concepts are the main attributes of the rating, which is used to report the size and strength of students' structures of knowledge. The paper describes the concept mapping scoring method, shows examples of students' work (including three-dimensional group concept maps) and discusses various aspects of this method.

Assessing Conceptual Understanding in Science through the Use of Two- and Three-Dimensional Concept Maps*

Michal S. Lomask and Joan Boykoff Baron, Connecticut State Department of Education[1]

BACKGROUND

Within both the science education and assessment communities, there continues to be a growing interest in the development of alternative approaches to assess students' understanding of important science concepts. Stimulated at least in part by a plethora of recent reports on what students in the United States should know and be able to do (AAAS, 1989; California State Department of Education, 1989; National Center for the Improvement of Education, see Raizen et.al. 1989, 1990; National Science Teachers Association, 1992) state departments of education, local school districts, and classroom teachers have been experimenting with different forms of alternative assessment. Alternative assessment uses a variety of formats, designed to elicit different skills and understandings. What they have in common is a movement away from norm referenced, standardized multiple-choice items to criterion referenced, open-ended, multiple-response types of items.

As part of a research and development effort, staff at the Connecticut Common Core of Learning Assessment Program (see Connecticut Board of Education, 1987) have developed three formats of assessment (called assessment shells) :

1) Performance tasks to assess students' scientific reasoning and experimentation (see Lomask, Baron, Greig, & Carlyon, 1992);

2) Integrated tasks to assess students' skills of data manipulation and science-based decision making, and

3) Essay-type tasks to assess levels of conceptual understanding.

The goal of this paper is to describe the essay-based assessment shell with its unique use of concept mapping as a tool for making meaning out of students' written responses.

UNDERSTANDING AND EXPLANATION

There are many definitions of understanding, some structural and some functional. In our project, we adopted the position that to understand something means to be able to explain it clearly to someone else. When learners build their understanding, they combine previous experiences with new ones to form a web of interconnected concepts. New experiences either add new concepts to existing structures of knowledge or they cause a shift to create a new structure. In a way, the individual learning process is similar to the general process of scientific endeavor in which new data either support current theories or cause a shift toward a new theory. The structures of knowledge (understanding) are defined by articulating the cognitive structure into words, sentences and logical explanations. The learner who understands a concept or a whole web of concepts should be able to explain these concepts and use his/her understanding in developing accurate descriptions and prediction of events that are related to the topic. But learning is not an isolated event and the understanding of concepts is a continuous process. Therefore, we assume that students will be found to have different levels of understanding. Some will have the basic, *nominal* level of scientific understanding (able to recognize terms only by their names, with loose association to central concepts). Others will acquire a *functional* level (able to memorize correct definitions of terms and concepts without understanding), or a *structural* level (able to construct relationships among concepts) or perhaps even a *multidimensional* level (able to apply and integrate knowledge of different concepts). (See BSCS, 1992). Believing in the tight relationship between understanding and explanation, the project aimed to uncover students' understanding based on their written explanations of natural phenomena and scientific processes.

ASSESSMENT FORMAT

The assessment was composed of various short items designed to elicit students' understanding of main concepts of science. Open-ended items that ask students to synthesize a written answer and to describe, predict and explain phenomena, are old veterans in the field of assessment. The main limitations on the use of essay-type items have been the low reliability of their scoring and the extended time that it take to score students' answers. The common method of analysis of essay-type answers is based on holistic scoring in which every answer is graded on a predetermined scale. The objectivity of the scoring process is affected by external factors such as writing style, grammar, etc.. This is a serious psychometric problem, but since students' own written explanation is the one most important indicator of their understanding, we decided to keep including essay-type items in our assessment and focus our efforts on increasing the reliability, efficiency and meaning of the scoring process.

ConnMap, Connecticut's concept mapping system, uses a concept-mapping approach, developed by the authors and refined in collaboration with more than a dozen high school science teachers, to analyze students' written responses. This work is a variant of the concept mapping approaches used by Champagne, Klopfer, DeSena & Squires, 1978; Hoz, Tomer & Tamir, 1990 and Novak & Gowin, 1984. The specific goal of our approach was to find a valid and reliable way to describe what students understand about important science topics based on their written essay responses. These descriptions could then be analyzed and reported in meaningful ways to students, teachers, curriculum and instructional designers, school administrators, local and state-level policy makers, and the general public.

CONNMAP - SCORING WITH CONCEPT MAPPING

The ConnMap method which was developed to score students' conceptual understanding via concept mapping, was used by high school science teachers, scientists and professional science educators. The complete scoring process consisted of three main stages:

I. Analysis of students' answers

II. Evaluation of students' answers

III. Evaluation of students' understanding

I. Analysis of students' answers

This stage consisted of three steps:

a) Drawing the expert's (teacher's, in this case) concept map for each specific item.

b) Translating the student's written response into a student-based concept map.

c) Comparing the student's map to the expert's map.

II. Evaluation of students' answers

The evaluation of students' answers was based on a review of the concept maps which were drawn in stage I. Based on these maps, two structural dimensions were identified:

1. **Size** of structure - Size (X) is determined from the *proportion* of the number of concepts described by the student to the number of concepts included in the expert's concept map. That is, the denominator of the fraction is the number of concepts in the expert's map and the numerator is the number of concepts in the student's map that are found in the expert's map. If the student includes additional related concepts in his/her map, whether they are correct or incorrect, the concept are added to the concept map, but they are not counted in the determination of size. The dimension of structure size has the following sub-levels:

Complete Structure	→	$80\% < X \leq 100\%$
Substantial Structure	→	$60\% < X \leq 80\%$
Partial Structure	→	$40\% < X \leq 60\%$
Small Structure	→	$20\% < X \leq 40\%$
Insignificant Structure	→	$0\% < X \leq 20\%$

2. **Strength** of structure - Strength (Y) is determined from the proportion of the number of *valid connections* made by the student to the number of connections expected based on only those concepts the student included. The strength of the student's map is related to the size of the

structure because the size determines the number of expected connections. For example, suppose a student mentions only three concepts - A, B & C - to an item that contains nine concepts in its expert map. If, in the expert map, A is connected to B and B is connected to C, we would expect a maximum of two connections. Therefore, in this case the expected denominator would be 2 and the numerator will be determined by the actual number of correct connections made. The dimension of structure strength has the following sub-levels:

Strong Structure → $66\% < Y \leq 100\%$
Medium Structure → $33\% < Y \leq 66\%$
Weak Structure → $0\% < Y \leq 33\%$

It is important to mention that the expert map has been constructed to serve as a flexible guide for the rater to use in interpreting a student's written essay; it is not intended to serve as a straitjacket. Therefore, we ask raters to accept alternative parallel wording on the part of the student for concepts and connections included in the expert map. For example, in the energy transformation item, we accept "stored energy", "potential energy" and "chemical energy" as parallel terms that define the same concept. Related concepts that are included in the student's answer but not in the expert map, are kept by drawing them into the student's map; however, they are not counted in the scoring process. In this way, rich and instructionally important information is kept without interfering with the reliability of the scoring decisions made by the raters.

III. Evaluation of students' understanding

Overall scoring of the response, which reflects the understanding of the student, was based on the combination of the size and strength dimensions using Table 1, below.

Table 1
Combination of Structural Dimensions to Reflect Level of Student Understanding

Size	*Strength*: Strong	Medium	Weak
Complete	5	4	3
Substantial	4	3	2
Partial	3	2	1
Small	2	1	1
Insignificant	1	1	1

The student's overall score is represented by the number in the cell that shows the intersection of the size and strength scores. Even though there are 15 cells, there are only 5 score categories which represent different levels of conceptual understanding, as shown in Table 2.

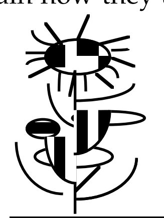
Table 2
Five Categories of Students' Performance Representing Level of Conceptual Understanding and Cognitive Structure

Cell Score	Level of Conceptual Understanding	Cognitive Structure
1	Novice	Nominal
2	Apprentice	Functional
3	Proficient	Structural
4	Accomplished	Multidimensional
5	Distinguished	Multidimensional

EXAMPLES OF SCORING

In this section we display five examples of students' work and how they were scored by the ConnMap system. One of the items in the biology section asked the students the following:

"Describe the possible forms of energy and types of materials which are involved in the process of a growing plant and explain how they are related to each other."



First, a group of five experienced high school teachers (biology teachers, in this specific example) working individually, composed an answer to the item, and then together constructed a concept map that represented the best description of their answers (see Figure 1.)

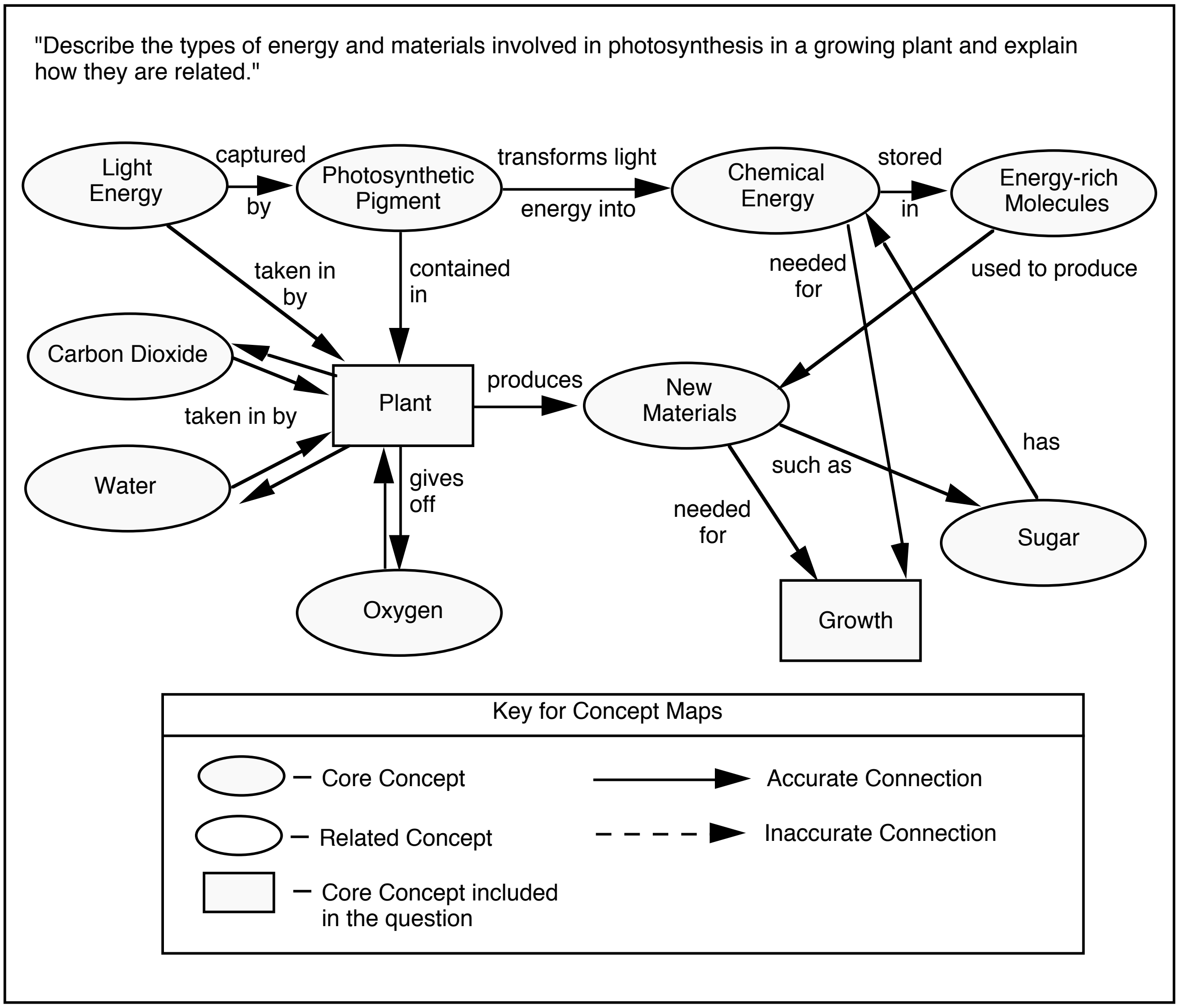


Figure 1. Expert concept map for "Growing Plant" item.

This expert concept map served as the goal construct for the teachers' later analysis of students' answers. Teachers then read the students' responses and constructed a map for each student.

The following five figures shows students' original written answers and their conversion into concept maps. Since the teachers used the expert map as a template, the terms used in the teachers' map can be represented by many parallel terms in the students' responses.

Based on the constructed students' maps, the teachers calculated the size and strength of the student's structure of knowledge. The overall score was determined by combining these two dimensions using the intersections on Table 1. The respective size, strength, and overall score for five students whose answers are illustrated in Figures 2-6 are summarized in Table 3.

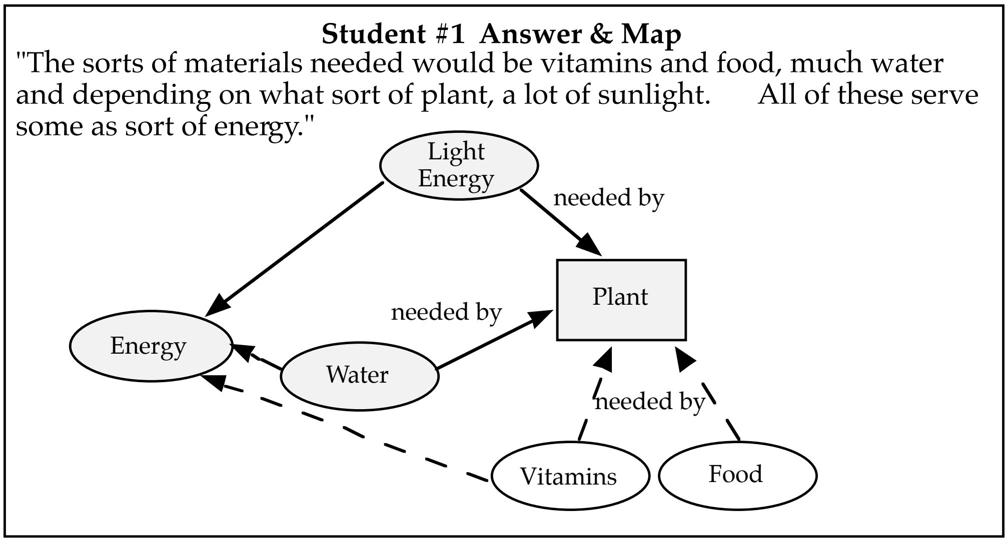


Figure 2. Example of a student response receiving a score of 1.

Figure 2 shows that five concepts were included in the student's answer. Three of the five concepts were included in the expert concept map and were therefore counted in determining the size. Two of the concepts were connected to the main concept in a valid, though shallow, way. The third and fourth (dashed line) connections show a common misconception, i.e. that plants use food and vitamins from an external source which serves as energy for the plant. This answer is at the novice level of understanding, typical of students with a nominal cognitive structure.

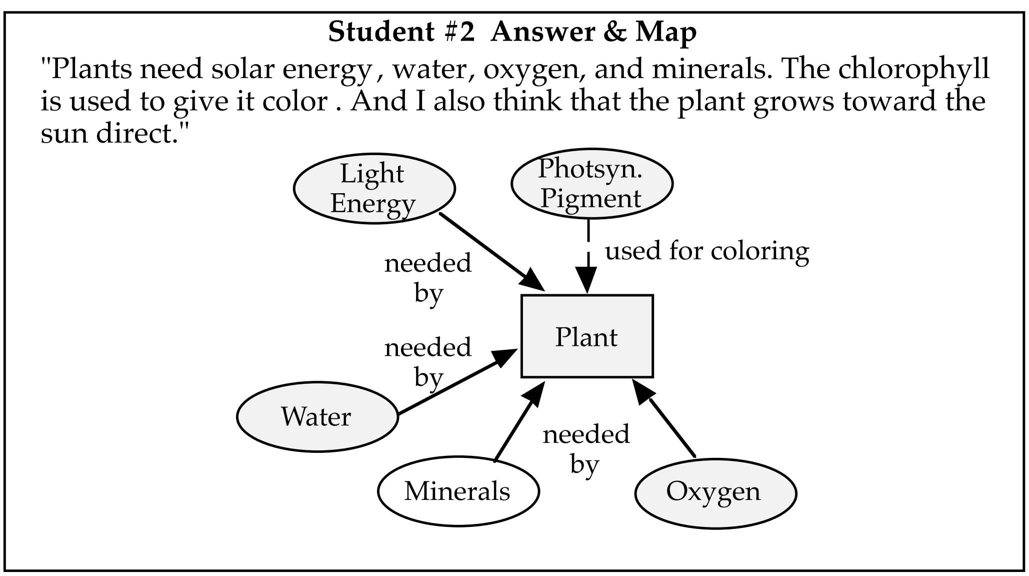


Figure 3. Example of a student response receiving a score of 2.

Figure 3 shows a typical answer for a level 2 student response. Two additional concepts are included (photosynthetic pigment/chlorophyll and oxygen) but their connection to the structure is incomplete and sometimes inaccurate. This answer is at the apprentice level of understanding typical of students with a functional cognitive structure.

Student #3 Answer & Map

"The growing plant has chlor oplasts that absorb light energy from the sun and use it for photosynthesis which provides the ener gy necessary for plant growth. The plant also needs carbon dioxide for this process."

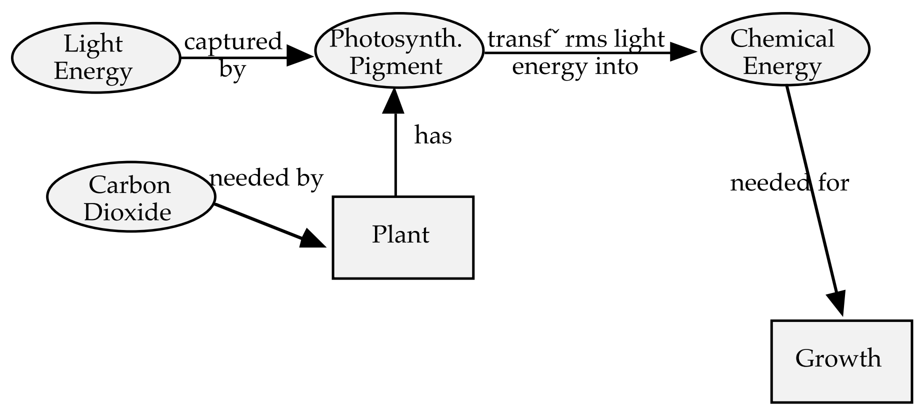


Figure 4. Example of a student response receiving a score of 3.

Figure 4 shows a typical level 3 response. Although the total number of core concepts still represents only partial size, the connections among the concepts are correct and create a reasonable explanation of energy transformation. This answer is at the proficient level of understanding typical of students with structural cognitive structures.

Student #4 Answer & Map

"The sun's energy is taken in by green chloroplasts in the light phase of photosynthesis. Carbon dioxide is taken in in the dark phase and is mixed with water and energy from the sun to produce glucose, which is the energy stored in the plant."

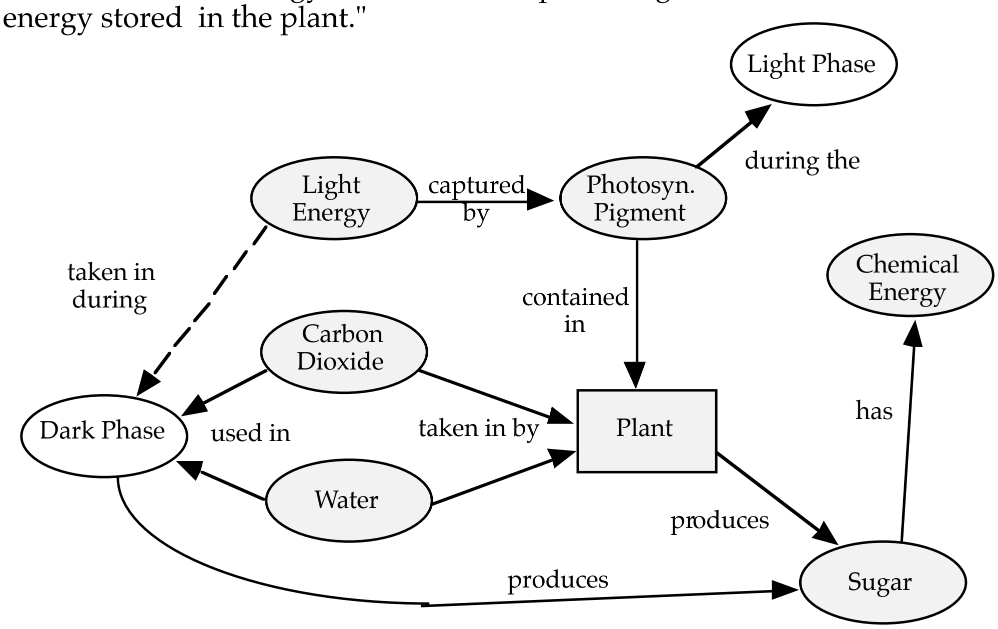


Figure 5. Example of a student response receiving a score of 4.

Figure 5 shows a typical answer for level 4. The number of concepts included in the answer is substantial and most of the connections are valid. This student included two related concepts (light and dark phase) that were not originally on the expert map. These concepts are certainly part of the conceptual structure under discussion and their omission from the expert map represents decisions made by the teachers regarding the level of detail expected from high school students. Although the answer shows a deeper understanding of the concepts involved, the relationship between energy transformation and materials exchange in the plant is not accurate. This answer is at the accomplished level of understanding, typical of students with low multi-dimensional cognitive structures.

Student #5 Answer & Map

"The main source of energy for growing plants comes from the sun. The plants use photosynthesis to harness this energy and construct ATP's (adenosine tri-phosphate) which is the basic source of energy in many organisms. They then use some energy harnessed from the "light" reactions in A TP's to do the "dark" r eactions in which a glucose molecule is constructed. This easily transportable molecule contains much potential energy and can be distibuted throughout the growing plant. Many other nutrients and substances ar e needed for these r eactions. Atoms like nitrogen, converted by bacteria, and water are taken in through the roots and vascullarly transported to the rest of the plant. Oxygen and carbon dioxide ar e taken in through openings called stomata in the leaves. These substances provide raw material for plant growth and a means for the plant to rid itself of waste."

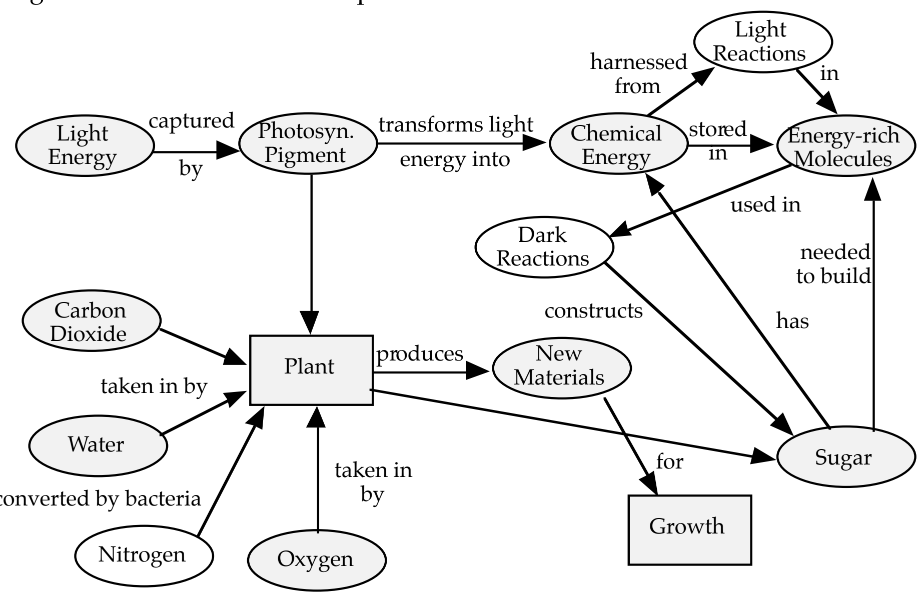


Figure 6. Example of a student response receiving a score of 5.

Figure 6 shows a high level (5) student's response. It includes all of the concepts on the expert map plus several more (light and dark reactions in photosynthesis, nitrogen fixation by bacteria, role of stomata, etc.) and all the connections are elaborated and valid. This answer is at the distinguished level of understanding, typical of students with high multi-dimensional cognitive structures

Table 3 summarizes the date for the five students, whose work was presented above.

Table 3

Summary of Dimensions Analysis for Five Different Student Responses

Dimension	Student #1	Student #2	Student #3	Student #4	Student #5
Size (concepts	2/9 Small	4/9 Partial	4/9 Partial	6/9 Substantial	9/9 Complete
Strength (connections)	2/3 Medium	3/5 Medium	5/5 Strong	9/9 Strong	15/17 Strong
Overall Score	1	2	3	4	5

Summary of ConnMap Scoring Method

- Analysis of students' answers is based on transformation of the written responses into visual concept maps and comparing them to an expert concept map.
- Two structural dimension of knowledge (size and strength) are identified and used for the scoring of students' answers.
- The use of concept mapping techniques enables the rater to keep a detailed visual representation of the student's conceptual understanding.

ANALYSIS OF A GROUP'S UNDERSTANDING

For program evaluation, curriculum revision, and instructional improvement it is important to know how a given group, as a whole, performed on each item. For this purpose, a three-dimensional concept map was constructed. In their analysis of students' work, the teachers used the expert concept map as a model; around it they shaped the students' maps (see Figures 2-6). When the maps from all of the students in a class were layered on each other, three-dimensional maps, representing the combined group response, were formed (see Figures 7 and 8). In the 3-D maps, the third dimension is proportional to the percentage of each concept and connection included in the answers of the group under study.

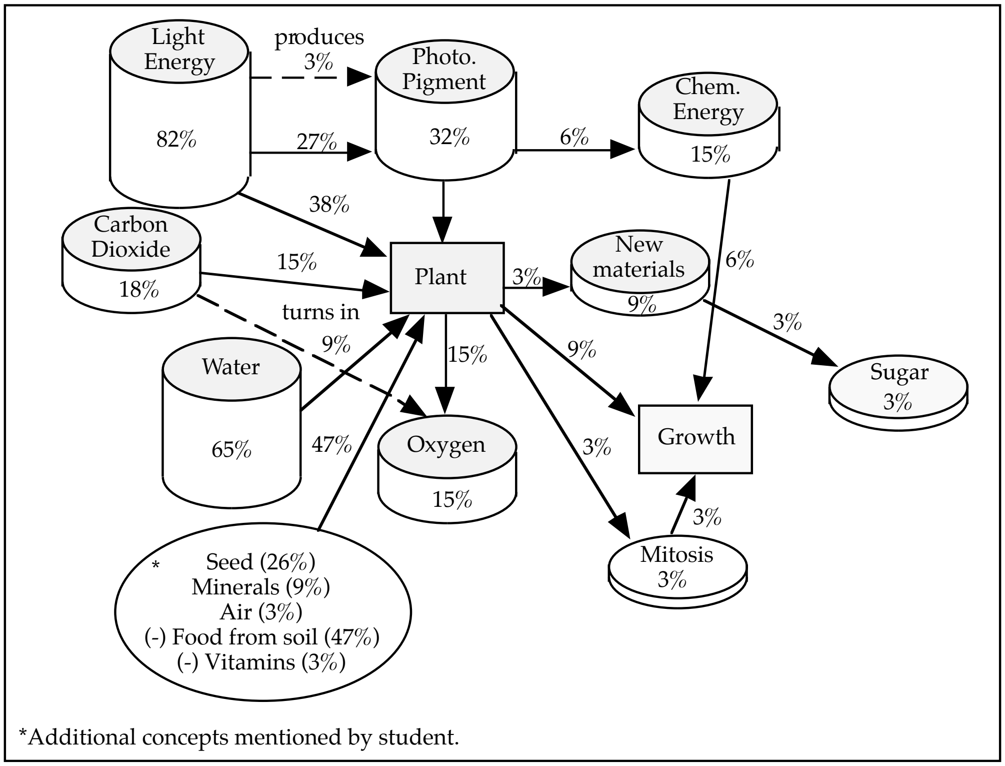


Figure 7. 3-D Map for First-Year Biology Classes (N=34)

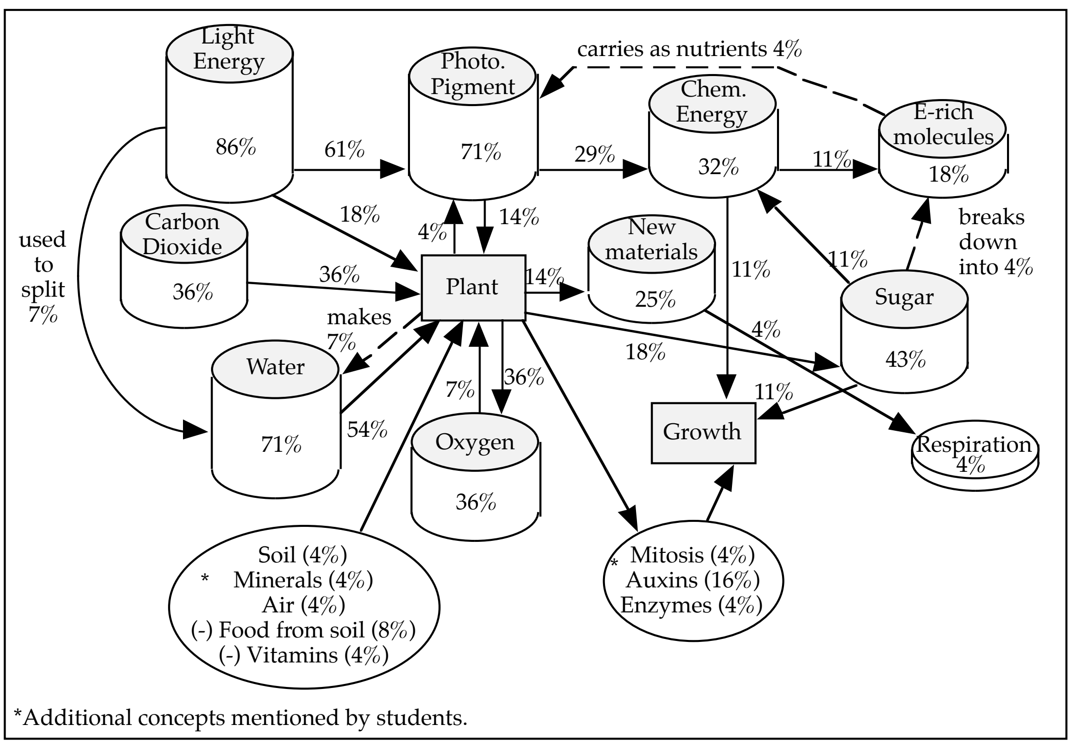


Figure 8. 3-D Map for AP Biology Classes (N=28)

Figure 7 shows a 3-D map that was constructed based on the answers of thirty-four first-year biology students and Figure 8 represents twenty-eight Advanced Placement biology students. Careful analysis of these maps shows that the AP students, as a group, have a bigger and stronger structure of knowledge than the first-year students. Many of the AP students mastered the idea of energy transformation as a basic characteristic of plant growth. Few of these students included the materials and how they are synthesized, a pattern even more true for the first-year biology class. We also see that almost half of the regular biology students (47%) have a common misconception, believing plants obtain food from the soil.

Another example of the rich information that is gained from the 3-D maps is shown through the results of a second item in the biology assessment. In this item, students were asked to answer the following question:

> If you were going to receive a blood transfusion, for what and why would you want the blood to be checked? Please explain as clearly as you can.

The teachers' concept map for this item is shown in Figure 9. The map contains two paths; one deals with the compatibility of blood types and the second with the prevention of infectious diseases. Teachers expected their students to understand the reaction between RBC-surface antigens and natural circulating antibodies as the basis for blood compatibility. They also expected their students to be able to distinguish infectious diseases which might be transferred through body fluids, like AIDS, hepatitis and syphilis from non-infectious diseases like diabetics or leukemia.

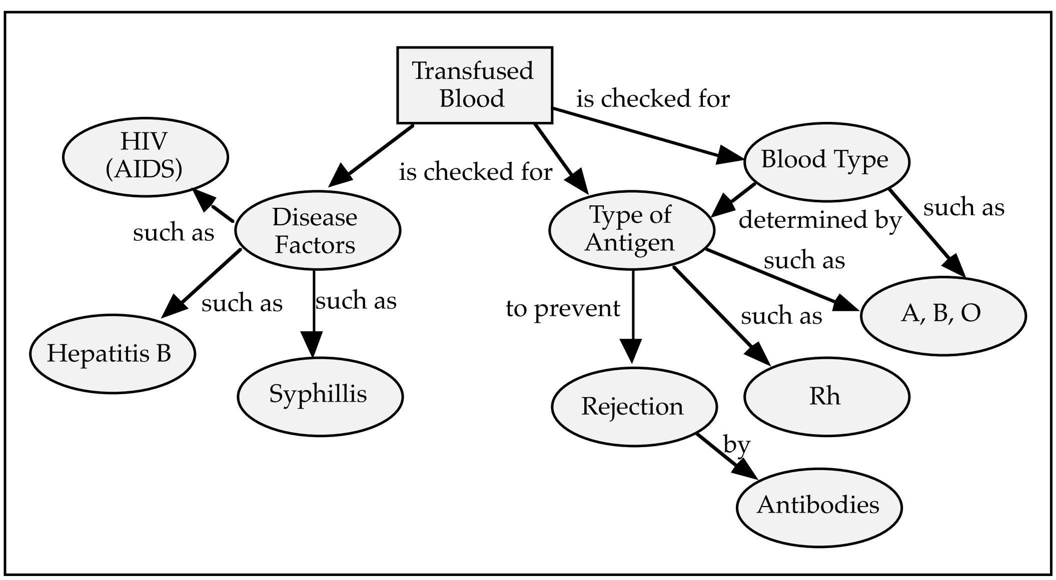


Figure 9. Expert concept map for "Blood Transfusion" item.

Figure 10 shows 3-D map, built upon first-year biology students' answers. The map reveals students' high awareness of the risk of contracting AIDS, compared to the low awareness of the risk of blood incompatibility.

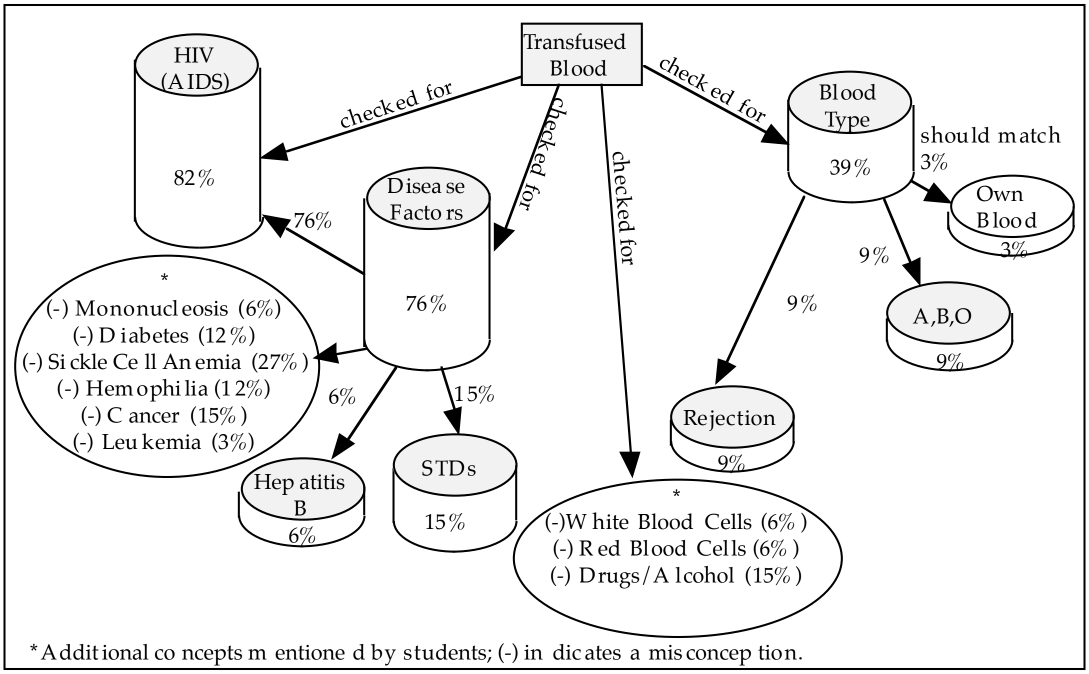


Figure 10. 3-D Map for First-Year Biology Students (N=33)

Figure 11 shows 3-D map, built upon AP biology student's answers. The map reveals a higher understanding of blood compatibility, but many errors about disease factors.

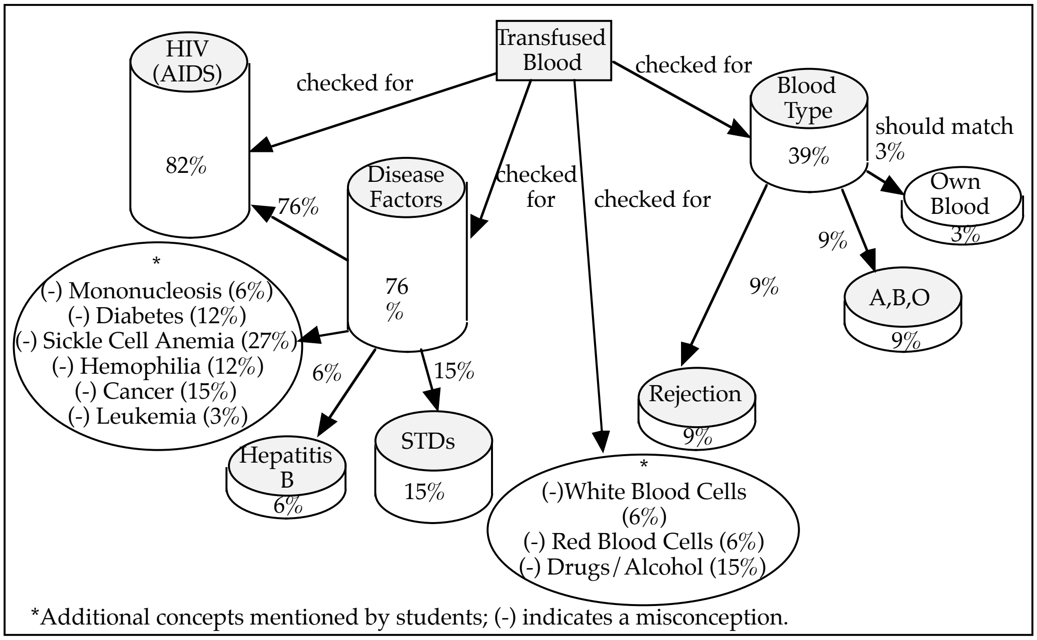


Figure 11. 3-D Map for AP Biology Classes (N=2)

Comparisons of the two 3-D maps created for the Blood Transfusion task show the following:

a. Most of the students are aware of the risk of contracting AIDS through contact with blood which contains the HIV.

b. About half of the students are aware of the need to match the blood types of donor and receiver.

c. Based on the frequency with which they were mentioned and comments made by the students, it seems that most of the students (more than 85%) believe the consequences of contracting AIDS are more dangerous than the consequences of receiving incompatible blood.

d. Many students don't know the difference between genetic and infectious diseases.

e. AP biology students have more elaborated knowledge about blood compatibility than first-year biology students. No significant differences were observed between the two groups regarding transfer of diseases. This suggests that either everyone learns about disease factors in the first year of biology or more likely, informal learning (e.g. media exposure, etc.) is affecting students' knowledge of this particular topic.

The main advantage of the 3-D maps is that they serve as tools for the evaluation of the cognitive structures of groups of students. They allow teachers to see at a glance which concepts and connections are included and which are omitted. They also allow for visual comparisons of the structures of different groups of students, a feature that might help the teachers to reshape learning materials and plan appropriate instructional units. Additional examples of 2-D and 3-D concept maps are provided in Appendix 1.

DISCUSSION

1. <u>Reliability.</u> Can different raters *reliably* score the student's response; i.e., will several raters score the same student's response consistently? In order to determine the reliability of the scores produced by the system, we asked four biology teachers to score the same students' work. The responses of 39 students to the "Growing Plant" item and 42 responses to the "Blood Transfusion" item were scored. Using the expert maps, the four teachers independently produced concept maps representing the students' responses. For each item, the following three aspects of the concept maps were considered:

1. the number of concepts on the expert map that were used by the student, (this is the numerator of the size proportion);
2. the number of correct connections between concepts that were used by the student, including both those on the expert map and those relevant concepts added by the students; (this is the numerator of the strength proportion);
3. the expected number of connections among all of the concepts mentioned in the essay; (this is the denominator of the strength proportion).

The data were analyzed using the approach of Generalizability theory (see Shavelson, Webb, and Rowley, 1989). In a G study, a variance component is computed for each source of variation, including systematic variance among the objects of measurement, and multiple error sources. For each of the three aspects of the concept maps (number of concepts, number of correct connections, and expected number of connections) a variance component due to each of the following sources was estimated:

1. variance due to differences among *students* (true variance),
2. variance due to systematic bias among *raters*,
3. variance due to all *other sources*, including interaction between students and raters.

Table 5 displays, for each aspect of the concept maps, the percentage of total variance due to each source.

Table 5
Results of Reliability Study of Concept Map Elements

	Growing Plant			Blood Transfusion		
Source of Variance	Concepts	Correct Connections	Expected Connections	Concepts	Correct Connections	Expected Connections
Student	84%	77%	81%	89%	84%	62%
Rater	0%	0%	1%	1%	1%	9%
Error	16%	23%	18%	10%	15%	29%

The results were encouraging, although some difficulties should be noted. As indicated by the relatively high percentage of variance due to differences among students, the teachers were most consistent in determining the number of concepts on the expert map used by the student (84% for the growing plant task and 89% for the blood transfusion task.) They were somewhat less consistent in determining the number of correct and expected connections. For the relatively complex concept map for the growing plant item, 23% of the variance in scoring the number of correct connections is attributable to error. For the blood transfusion item, 9% of the variance in scoring the expected number of connections is due to systematic differences among raters, and 29% to error. One problem, especially apparent in the blood transfusion item, was that some teachers noted additional concepts on their maps included in the students' answers, while others did not. However, it seems that, with *appropriate and sufficient training,* raters can reach a higher level of agreement on the scoring elements, thus supporting the reliability of scores obtained by the ConnMap method.

2. Feasibility. Is this scoring methodology feasible? Are teachers comfortable using concept maps to score students' responses? What impact might using the ConnMap scoring procedure have on teachers? Four Connecticut biology teachers, who had been working with us on refining the concept mapping procedure, received about an hour of training on the latest version of the concept mapping procedure. During this time, they scored several papers and then discussed how they had reached their scores. When agreement on scoring was reached, the scorers then worked independently, scoring approximately 80 papers in about three-hours. These four teachers reported feeling comfortable with the concept-mapping procedure. Before reporting further on teachers' comfort with the procedure, we will need to train several groups of teachers who have had no previous contact with the item development process.

3. Time. How long does the scoring take? At the current time, using a pencil and paper mapping procedure, it takes teachers approximately two to three minutes per student to develop the concept map and compute, by hand or with a calculator, the size and strength of the student's map. We believe that this time can be reduced by using a software program that we hope to develop.

4. Usefulness. Is the information obtained from using concept maps to score students' responses **useful** to students? Do teachers find the individual profiles of students useful for working with students? Do teachers find the class profiles useful for instructional planning? Are the state-wide profiles useful for policy makers in determining whether programs are successful? To answer these questions, we will need to look separately at the reporting elements for each constituency. Will the students find the pictorial visualizations of their partial concept maps (see Figures 2-6) helpful in constructing deeper understanding of scientific concepts? The

literature on the importance of accessing prior knowledge suggests promising results in this regard. Whether the teachers will find both the individual and classroom 3-D maps (see Figures 7 and 10) useful in working with individual students and revising their curriculum and instruction is still an open question. Similarly, whether district and state-level policy makers will find the score distributions in Table 3 and Figure 8 useful is still unanswered. They should permit policy makers to monitor student growth in much the same way as they monitor changes in the distributions of holistic writing scores. We would anticipate that our 3-D maps would allow science education researchers and program designers to monitor the percentages of students abandoning naive or non canonical science conceptions.

5. Impact on teaching. Related to the utility of the results in the utility of the process itself. Another aspect to this question is whether the involvement of teachers in developing and/or using the concept maps serves as professional development for teachers, *by helping them to develop additional content expertise.* Preliminary evidence indicates that the participating teachers feel that the experience has helped them to better understand the content being assessed and that as a result they can teach the concepts more effectively. Hopefully, these beliefs will be corroborated in studies of the consequential validity of this assessment.

6. Consequences. The process of developing the expert concept map and then mapping student responses was also very useful to the assessment developers in evaluating the success of the open-ended questions. The concept maps were particularly helpful in determining if the question truly tapped into the intended scientific concepts and connections and in determining if the questions were clear enough to elicit the kind of responses that would provide evidence of students' understanding. As a result of this scoring procedure many questions were reworded and repiloted. The list of unintended consequences will undoubtedly grow once this procedure is implemented.

7. The Use of Computer Technology to Generate and Score Student's Concept Maps. In the interests of both accuracy and efficiency, we are currently taking steps to develop software to allow teachers to enter the information from students' responses directly onto a computer. This will replace the use of the current pencil and paper technique in which teachers record the student's concepts and connections on the shell of a prepared expert map. It will also eliminate the counting and calculations of the percentages of concepts and connections currently required by our method. However, the computer will not eliminate the need for experienced high school science teachers who will continue to be vital to the scoring process. They will make and enter judgments about the presence of the concepts and the connections in the student's response.

Specifically, they will determine whether the concepts that the student mentions are parallel to and consistent with the expert's map. They will also determine whether the connections that the students make (or imply) are accurate. These judgments will be entered directly onto the expert concept map which will have been stored in the computer. When the data have been entered, the computer will calculate the size, strength and overall score for each student and group of students. The computer will also generate the student's pictorial map, 3-D group maps, and matrices and histograms of score distributions.

Our work will build upon the work that has already been done to allow students to create their concept maps directly on a computer. Fisher (1990) provides a review of several types of software for concept mapping and semantic mapping, and Baker, Niemi, Novak & Herl (1991) provide an application using a version of Hypertext.

CONCLUSION

We see several advantages in using the ConnMap technique to describe, interpret and report students' understanding of important scientific concepts. The assessment of the size and strength of students' conceptual understanding should remove some of the subjectivity inherent in holistic scoring approaches. We therefore hope that the scores will be perceived by students, teachers, school administrators, parents, and policy makers as more reliable and trustworthy. This, in turn, will lend support to the use of essay responses in alternative models of student learning and assessment.

We further hope that by using essay-type responses we are encouraging a depth of understanding that allows students (a) to determine when various concepts are relevant to a problem; (b) to tie concepts together to tell a whole story, and (c) to develop their writing skills. We hope that by involving science teachers in all phases of the scoring, they will (a) enhance their own professional knowledge and (b) be able to use these kinds of essay questions and scoring procedures in their own classroom-based assessments.

---

[1] The stimulus for this work was the Connecticut Common Core of Learning Document (Connecticut State Board of Education, 1987) which demanded new forms of assessment to ascertain the extent to which students in Connecticut were attaining the desirable knowledge, skills, and dispositions set forth by a representative group of Connecticut's educational and business communities. The research was funded by The Connecticut State Department of Education and a grant from The National Science Foundation (SPA-8954692). We thank the dedicated science teachers and their students who worked closely with us to develop these concept maps. An earlier version of this paper was presented at the National Association for Research in Science Teaching (NARST, 1992) and we are grateful to Audrey Champagne and

Joseph Novak for their helpful comments in their role as discussants. A longer version of this paper will appear in Assessment of Conceptual Understanding (Lomask, Baron & Greig) in Assessment as an opportunity to learn: The Connecticut Common Core of Learning Assessment in high school science and mathematics (Edited by J. B. Baron, in preparation).

REFERENCES

American Association for the Advancement of Science. (1989). Science for all Americans: A Project 2061 report on literacy goals in science, mathematics, and technology. Washington, DC: Author.

Baker, E.L., Niemi, D., Novak, J. & Herl, H. (1991). Hypertext as a strategy for teaching and assessing knowledge representation. Paper presented at NATO Advanced Research Workshop on "Instructional Design Models for Computer-based Learning Environments," July 1-4, Enchede, The Netherlands.

Biological Science Curriculum Studies (1992). Developing biological literacy. Colorado Springs, CO. Author.

California State Department of Education (1989) Science curriculum framework. Sacramento: CA: Author.

Champagne, A. B., Klopfer, L.E., DeSena, A. T. and Squires, D. A. (1978). Content structure in science instructional materials and knowledge structure in students' memories. (Monograph 22). Pittsburgh: Learning Research & Development Center.

Connecticut State Board of Education (1987). Connecticut Common Core of Learning. Hartford, CT: Author.

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

Hoz, R., Tomer, Y. & Tamir, P. (1990). The relations between disciplinary and pedagogical knowledge and the length of teaching experience of biology and geography teachers. Journal of Research in Science Teaching, 27, 973-985.

Lomask, M., Baron, J.B., Greig, J., & Carlyon, E. (1992) What do our students know? Assessing students' ability to think and act like scientists through performance assessment. Paper presented at the National Science Teachers Association Annual Meeting. Boston, Mass. March 26.

National Science Teachers Association (1992) Scope, sequence and coordination.

Novak, J. D. & Gowin, D. B. (1984). Learning how to learn, Cambridge University Press.

Raizen, S. A., Baron, J. B., Champagne, A. B., Haertel, E., Mullis I. V. S., & Oakes, J. (1989). Assessment in elementary school science. Washington, DC: National Center for Improving Science Education.

Raizen, S. A., Baron, J. B., Champagne, A. B., Haertel, E., Mullis I. V. S., & Oakes, J. (1990). Assessment in science education: The middle years.. Andover, MA: The Network.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989) Generalizability theory. American Psychologist, 44., 922-932.

APPENDIX 1

Additional Examples of 2-D and 3-D Concept Maps:

1.1 Shining More Light (physics)

1.2 Burning Wood (chemistry)

1.3 It's Elementary (chemistry)

1.4 Day and Night (earth science)

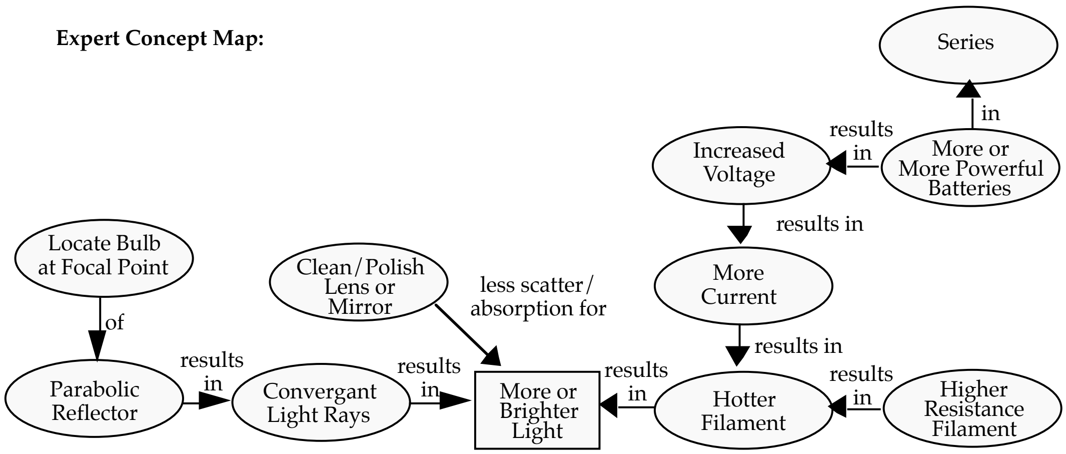
APPENDIX 1.1

Shining More Light: Task with Expert and 3-D Map

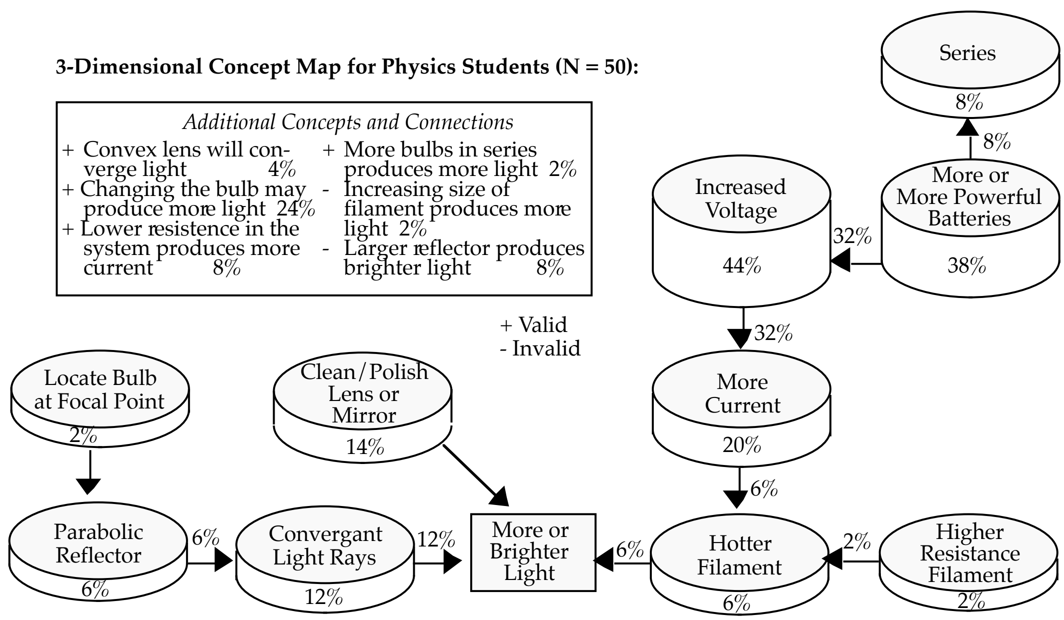
Shining More Light

How could you modify a flashlight to make it shine brighter? Explain fully why your modification would produce more or brighter light.

Expert Concept Map:



3-Dimensional Concept Map for Physics Students (N = 50):



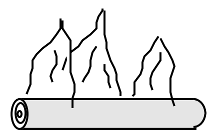
APPENDIX 1.2

Burning Wood: Task with Expert and 3-D Map

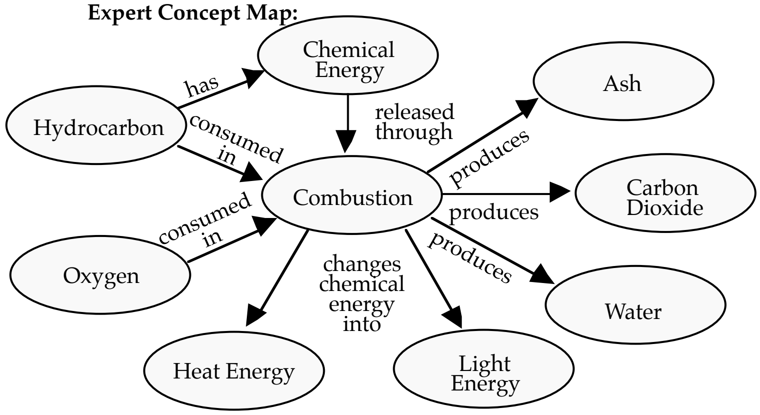
Energized Objects - Burning Wood

For the following situation, describe the possible forms of energy and types of materials involved and explain how they are related.

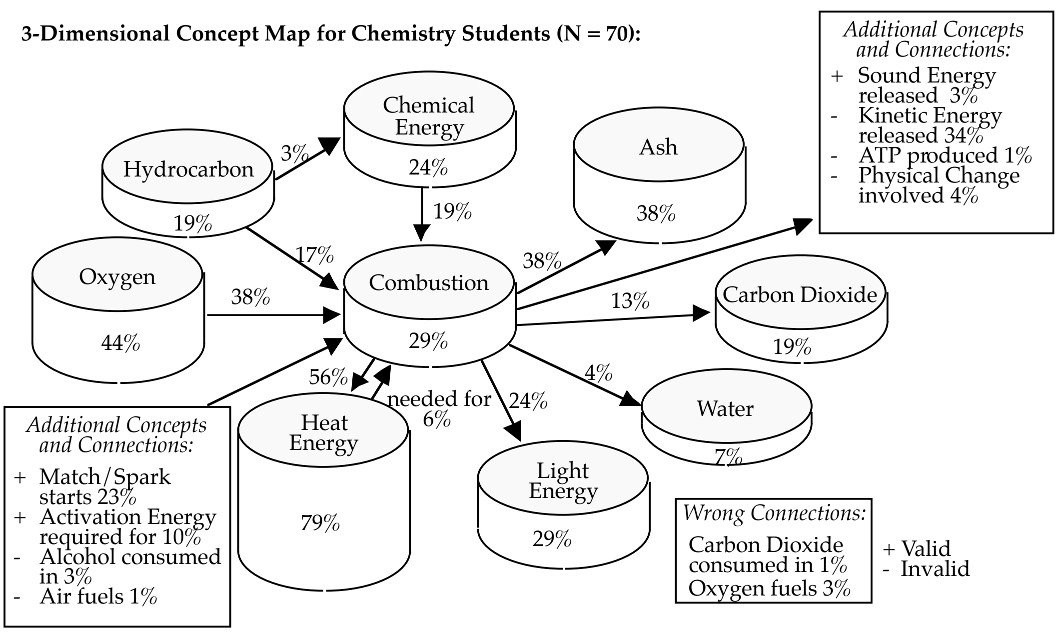
Burning wood:



Expert Concept Map:



3-Dimensional Concept Map for Chemistry Students (N = 70):



APPENDIX 1.3

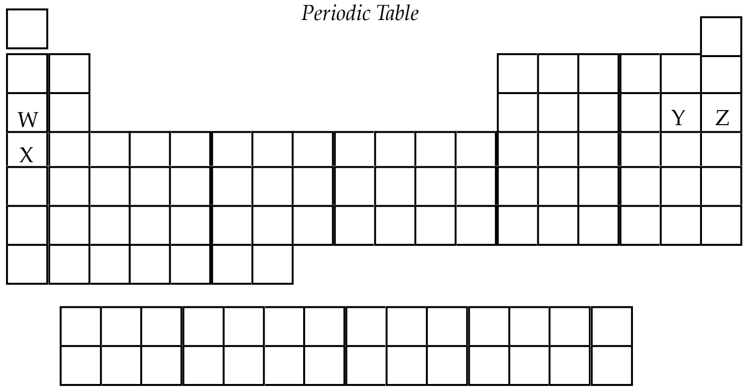
It's Elementary: Task with Expert Map

It's Elementary

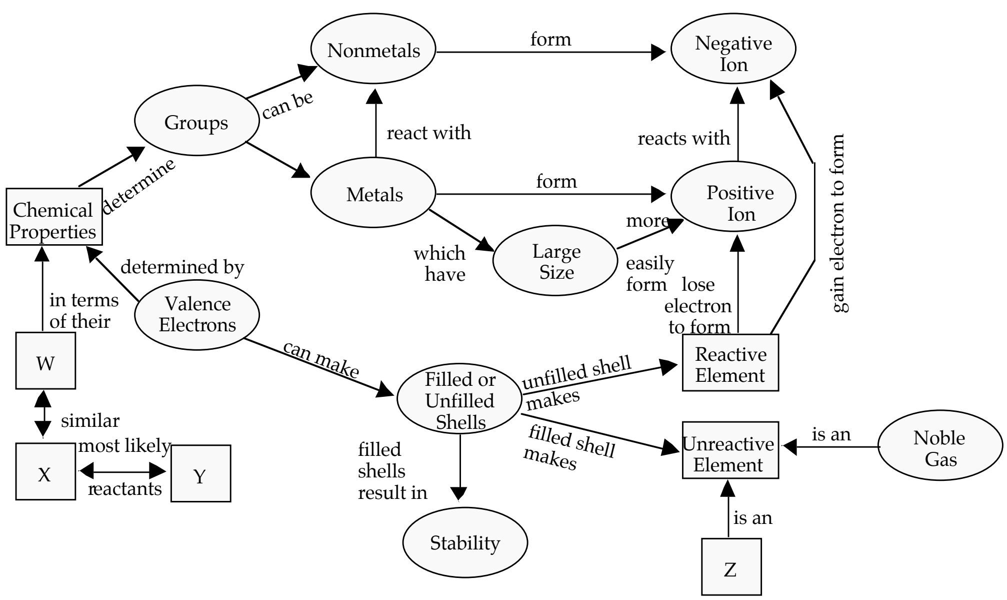
Use the periodic table below to answer the following questions:

1. W,X,Y, and Z are elements in this periodic table. Which two elements are the most similar in terms of their chemical properties? Explain fully why you think so.
2. Which of these four elements is the least reactive? Explain fully why you think so.
3. Which two elements are most likely to react with each other to form a compound? Explain fully.

Periodic Table



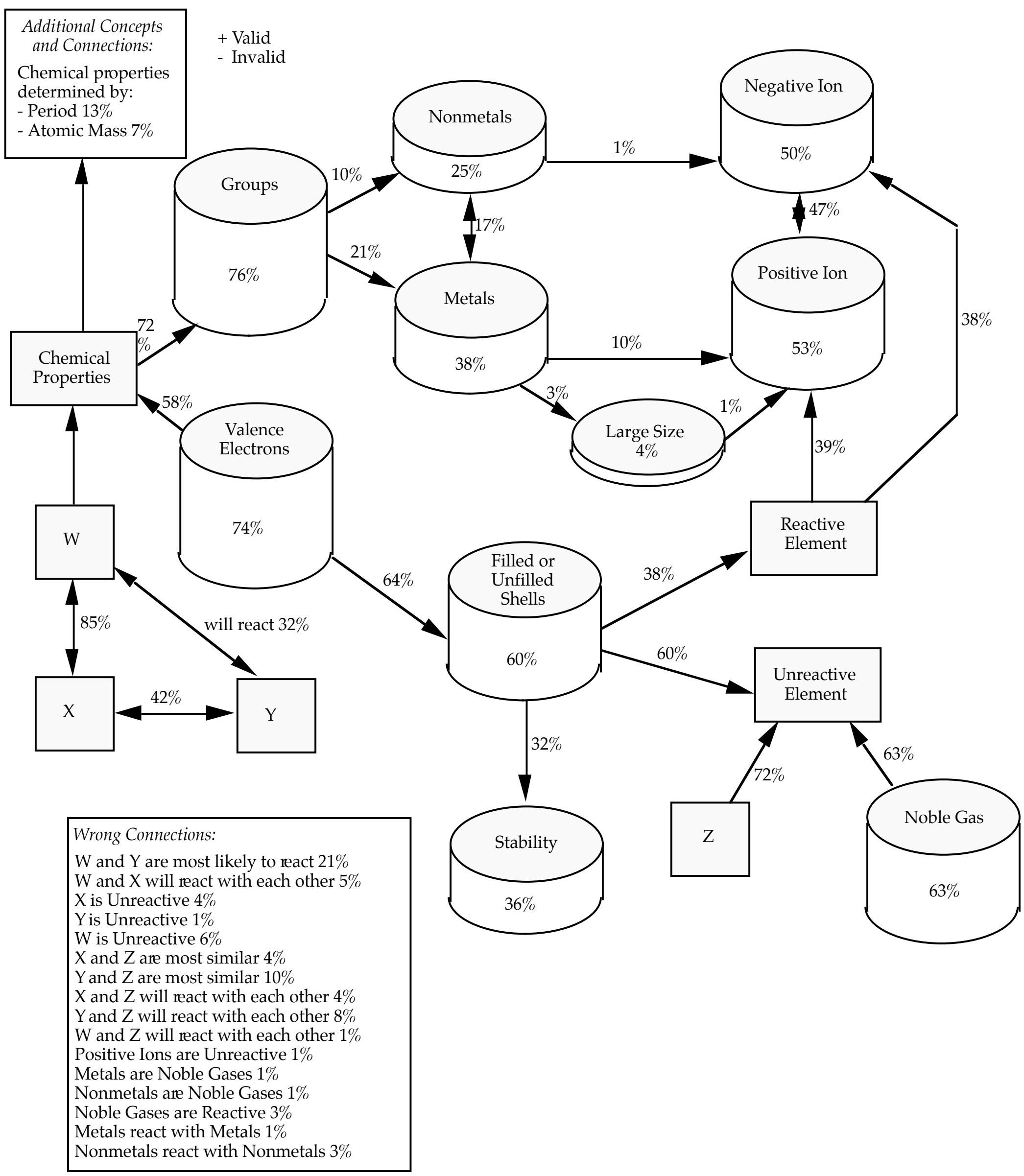
Expert Concept Map:



APPENDIX 1.3 Continued

It's Elementary: 3-D Map

3-Dimensional Concept Map for Chemistry Students (N = 72)



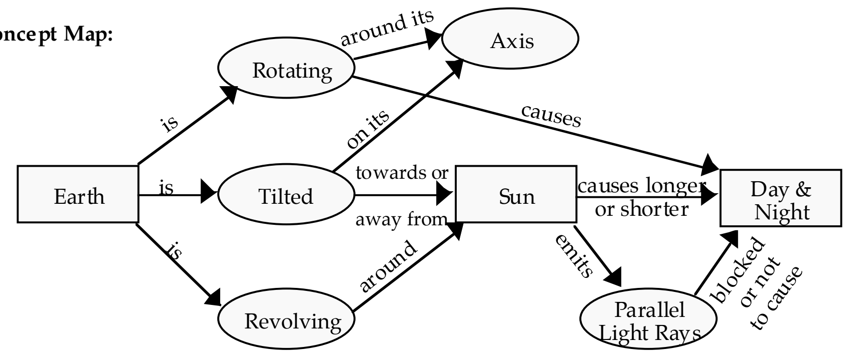
APPENDIX 1.4

Day and Night: Task with Expert and 3-D Map

Day and Night

We are all used to having daylight and night time every d ay. However, in the far north, above the arctic circle, there are days when the sun never rises and days when the sun never sets. Explain how this is possible. You may include a diagram if you wish. (If you include a diagram, be sure to explain it fully.)

Expert Concept Map:



3-Dimensional Concept Map for Earth Science Students (N = 49)