Concept Maps: Theory, Methodology, Technology
Proc. of the First Int. Conference on Concept Mapping
A. J. Cañas, J. D. Novak, F. M. González, Eds.
Pamplona, Spain 2004

PROMOTING DEEP APPROACHES TO LEARNING IN SCIENCE WITH SEQUENTIAL ASSESSMENTS

W. Bradley Kincaid
Mesa Community College, USA
kincaid@mail.mc.maricopa.edu, www.mc.maricopa.edu

Abstract. Scientific literacy requires deep learning of science concepts. A suite of sequentially dependent assessments, which includes online chapter quizzes, concept mapping, and essay exams, was employed to promote deep learning. These assessments were evaluated in a pilot study that included survey instruments for each assessment type and for the overall course as well as open-ended questions for each assessment. Analysis of the surveys and comments resulted in several hypotheses to be tested with subsequently collected data. Hypotheses from this pilot study include: 1) sequentially dependent assessments promote deep learning in science 2) adequate support and acceptance of concept mapping and essay exams affect student approaches to learning 3) concept mapping is effective for construction of knowledge and for developing the ability to communicate that knowledge 4) assessments that promote surface learning are a prerequisite for concept construction and communication.

1 Introduction

The problem of scientific literacy has been the subject of literally hundreds of reports over many decades (e.g., Educational Policies Commission, 1966). Science literacy is not simply measured by knowledge of science terms, nor is it necessary for all citizens to be able to perform as competent scientists (Shamos, 1995). Rather, what all citizens need is an understanding of our knowledge about nature, an ability to use the concepts of science as they show up in everyday life, and an awareness and appreciation of the process of science. Over the years, we have made progress in transforming the way we teach science to improve science literacy (e.g., Lawson, 1988). Regardless of our progress, much still remains to be done. Our students may have improved their achievement on objective tests, but, when asked to explain fundamental concepts, they fail. When asked to apply concepts in new situations, they fail. And when asked to make connections between related but separately discussed concepts, they fail. What seems to be happening is that students are failing to develop deep understanding of the concepts they are learning.

Recently, the National Research Council published a book linking research on *How People Learn* to classroom practice (Bradsford et al. 2000). Key findings of this synthesis are that 1) student preconceptions affect learning, 2) deep learning requires students to develop a conceptual framework for knowledge about a subject, and 3) a metacognitive approach to learning can help students achieve their learning goals. This excellent review concludes with implications for teaching to implement their key findings. Comparison of our practices in science education to these recommendations suggests that inquiry instruction provides an excellent structure for dealing with student preconceptions. However, science education does less well in helping students structure their conceptual knowledge and in promoting student monitoring of their progress toward learning goals. Much of what happens in the science classroom seems to guide students toward a surface approach to learning (Moore 1996, Novak, 2003). In fact, Halpern and Hakel (2002) assert that “it would be difficult to design an educational model that is more at odds with current research on human cognition than the one used in most colleges and universities.”

2 Alternative Assessment to Promote Deep Conceptual Learning

How can we motivate our students to achieve deep learning? Among other factors, assessment can have a significant impact on learning. Students seem to adapt their learning approach to the assessment expectations built into our courses. Thus, surface level assessment promotes surface level learning. In course design, our objectives must be embedded in assessment tasks. If for example course objectives focus on the big questions of biology, assessments must also focus on the big questions of biology. Our students should achieve understanding of those concepts and be able to communicate that understanding.

Three years ago, I undertook a redesign of my introductory biology course that primarily focused on matching my assessments to my course objectives. My resulting course design can be described in a concept map based on the sequentially dependent components of surface learning and concept construction to promote

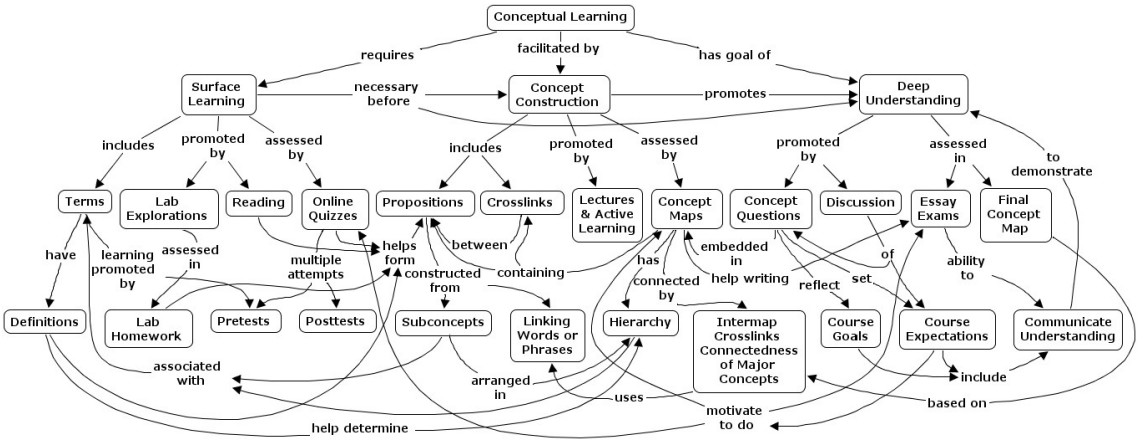

Figure 1. Concept map of sequentially dependent components of learning with assessments to promote deep conceptual learning.

deep understanding (Figure 1). Sequential assessments including online chapter quizzes, concept mapping, and essay exams are central to my educational model for promoting deep conceptual learning.

My assessment suite begins with online quizzes delivered in an online course management package. They were intended to reinforce reading and ensure surface learning. Each quiz comprised 20-40 questions for each of 9 chapters covered in the course. The focus was mastery of the vocabulary of the chapters. Students were given a list of terms for each chapter that they were expected to define and learn. I allowed students to have multiple attempts at taking these quizzes and gave them the highest score they received.

Concept mapping was employed to promote connections among the facts and theories associated with each biological concept (Novak, 2003). Concept mapping is a visual tool for construction of knowledge that can also be employed as an assessment tool. Students were required to use the term lists as the basis for a concept map of each chapter. In addition, I devised a final exam consisting solely of a course-level concept map wherein points were awarded only for valid crosslinks between the branches of a map that consisted roughly of the ten major concepts studied over the entire semester. Concept mapping can be especially valuable because of its ability to depict and promote learning of interrelations of concepts. I believe these crosslinks are the beginnings of deep understanding.

Finally, I employed essay exams as a final assessment of my student's deep understanding of the concepts covered in the course. I believe that writing about a concept is best test of deep understanding. Therefore, I wrote a set of 9-12 broad questions for each of my midterm essay exams, which were given to the students in advance. These were guided essay questions in that they prompted students to cover specific areas in their essays. For each of my three midterm exams, students were given four of the questions from which they could choose three on which to write their essays in class.

3 Student Evaluation of Assessments: Methods

To evaluate my alternative assessments, surveys were administered anonymously with our course management software. Survey questions included Likert-scaled items on 4 areas: overall course (26 questions), online quizzes (17), concept mapping (17), and essay exams (23). In addition, students were presented with an opportunity for open-ended comment on each assessment type. The survey was administered during the last week of classes.

Survey data were analyzed using principle components analysis of each survey area separately with varimax rotation to reduce the dimensionality of the data. Thus, a few new composite variables were produced for each assessment type, which retained most of the variation present in the original variables. These new variables can be interpreted by their correlations with the original questions and are independent variables with mean = 0 and variance = 1. Student comments were analyzed by scanning the responses to identify response categories. Then, all comments were distributed into one or more of the categories. Representative comments were excerpted, and the number of similar comments received was tallied. These qualitative data were used to substantiate the components of variation extracted from the survey data and the interrelations among student evaluations of the assessments and the overall course. Finally, simple correlations among factors (i.e., principle

components) from each of the survey instruments (online quizzes, concept mapping, essay exams and the overall course) were used to explore student perspectives of how assessments relate to aspects of the overall course and to each other. Data analyzed in this pilot study were from a single class during the fall of 2001. At high point, this class included 30 students of which 23 students completed the class and contributed to the data for this pilot study.

4 Student Evaluation of Assessments: Results of a Pilot Study

4.1 Online Quiz Components of Variation

Student comments on the online quizzes overwhelmingly support the idea that this assessment helped them learn. Fifty percent of the students gave extensive comments supporting this idea. However, a significant number also felt that they were too time-consuming. A few also felt that they helped their grades, but a third of the students thought that multiple attempts made the quizzes too easy. Analysis of the online quiz survey data produced three principle components (or factors) that accounted for 70% of the variation in the original 17 questions (Figures 2 and 3). The first factor, which alone accounted for 30% of the variation, was positively correlated with survey questions that asked about the effect of the quizzes on learning. Hence, it was interpreted as a general instructional value factor, which is supported by positive comments that the quizzes helped them learn, helped their grade, and prepared them for the essay exams. The second component was positively correlated with questions about wanting more quizzes and whether they were helpful and negatively correlated with questions about how time-consuming the quizzes were and whether they wanted fewer quizzes. This interpretation corresponds with comments that the quizzes were helpful but that they are too time-consuming and too long. The third factor was positively correlated with questions asking about feedback and fairness and seems to correspond with the comments that these quizzes were helpful (Figures 2 and 3).

4.2 Concept Mapping Components of Variation

The most frequent student comment about concept mapping was that it promoted understanding of the biology concepts covered in the course. In addition, nearly a third of the students felt the maps helped them do well on the essay exams. However, negative comments indicated that concept mapping is too time-consuming (32%) or too difficult to master and that inability to be successful at mapping frustrated them and may have impeded their understanding of biology concepts. Principle components analysis of the concept mapping survey data extracted three factors together accounting for 67% of the variation in the original 17 survey questions (Figures 2 and 3). The first factor was positively correlated with survey questions about whether this assessment was goal-oriented, required thought, and promoted understanding. Factor 2 was positively correlated with statements that maps were to time-consuming and too frequent and negatively correlated with wanting more concept mapping. The third factor was interpreted as reflecting positive aspects of concept mapping including that the effort required was reasonable, that they were challenging and that they were carefully chosen. Factors one and three seem to correspond to student comments that concept mapping promoted understanding and that they helped them do well on exams, whereas factor two seems to be most related to the negative comments about them being too time-consuming and frustrating (Figures 2 and 3).

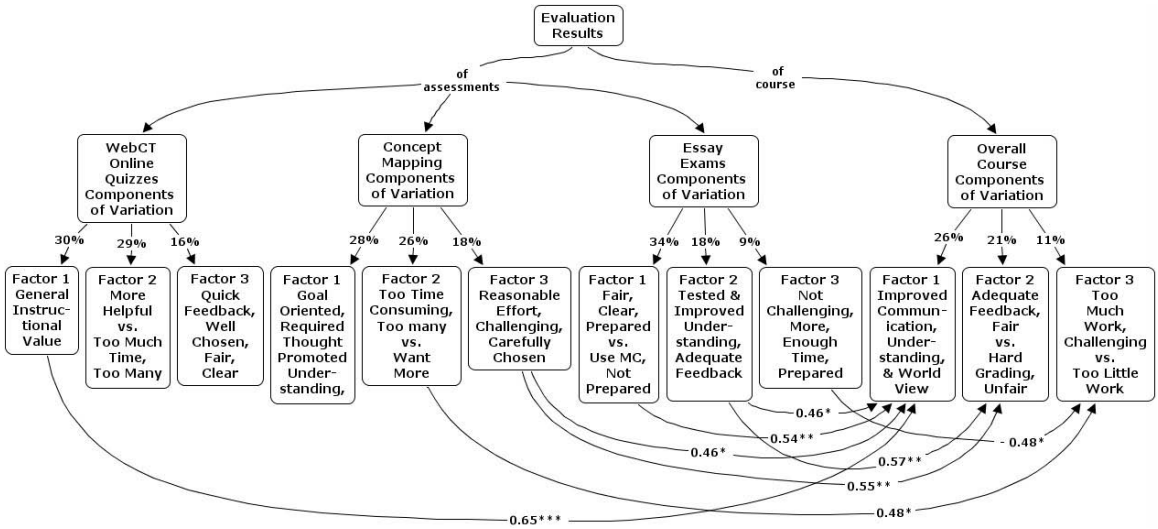

Figure 2. Concept map of evaluation results showing principle components for each survey instrument. Interpretations of each component are based on their correlations with the original survey variables. Percentages indicate variance explained by each factor. Intercorrelations of assessment components with overall course components are shown by the crosslinks and correlation coefficients. Correlation coefficients are shown for significant correlations between factors. Asterisks indicate significance of correlations: * = 0.05, ** = 0.01 and *** = 0.001. Eight significant correlations out of 27 considered are greater than expected by chance alone.

4.3 Essay Exam Components of Variation

Student comments presented a bipolar view of essay exams. The majority of comments were positive about this assessment type. Many students stated that essays either required or promoted understanding of biology concepts and that they were fair and effective assessments of their learning. Many also just liked this exam format over multiple choice exams. However, a significant number of students did not like the essay exams and suggested use of multiple choice exams instead. A few felt that essays were not an accurate reflection of their understanding and that they penalized students who did not write well. A few also felt that the class did not prepare them well for the essay exams. Principle components analysis of the essay exam data extracted three components that together explained 61% of the variation in the original 23 survey questions (Figures 2 and 3). The first factor was bipolar and alone explained 34% of the variation. Positive correlations with factor 1 were that the exams were fair and clear and that students were prepared for them. This factor was negatively correlated with the opinion that multiple choice should be used and that students were not prepared. Hence, this factor reflects a clear preference for the essay exams versus a preference for multiple choice exams. Essay exam factor two was positively correlated with the belief that the essay exams tested and improved understanding and that students were provided with adequate feedback. The third factor, although explaining a relatively small portion of the variation, was positively correlated with the belief that the exams were not challenging, that more exams should have been given, and that they were prepared. These factors correspond well with the comments about essay exams. In particular, factor 1 reflects the strong dichotomy between preference for essay exams and preference for multiple choice exams (Figures 2 and 3).

4.4 Overall Course Components of Variation

Principle components analysis of overall course survey produced three factors collectively explaining 58% of the variation in the original 26 survey questions (Figure 2). Overall course component one was positively correlated with student opinions that the course improved their communication skills and understanding of biology and that it broadened their worldview. Thus, this component is clearly aligned with my stated course goals of promoting understanding of biology and science process and affecting their worldview to include the interrelatedness of all life. Overall course factor two was positively correlated with the opinion that adequate feedback was provided and that they were fairly graded and negatively correlated with the belief that the course was hard and unfair. Course factor three clearly reflects a contrast in opinion that the course was too much work versus that it was too little work (Figure 2). Hence, we might consider factors one to three as the course goals, fairness and workload factors, respectively. Comments regarding the overall course were not explicitly requested, but the dimensions of variation extracted from the overall course survey can be seen in the comments for each of the assessment types.

4.5 Relationships of Assessment Factors with Overall Course Factors

To assess the relationship of assessment types with the overall course evaluation, factor scores for each of the assessment analyses were correlated with the overall course factor scores. The rationale for the approach is that the overall course evaluation should be dependent to a large degree on the assessments employed. Examination of these relationships might then reveal aspects the assessments that are associated with deep approaches to learning. Correlation of the assessment factors with overall course factor 1 suggests that some aspects of each of the assessment types contributed to student beliefs that they had improved communication, understanding and worldview (Figure 2). Of these, two essay exam factors produced significant correlations including the fairness/preparedness factor and the tested/improved understanding factors. Concept mapping (factor 3) and online quizzes (factor 2) were also positively correlated with overall course factor 1 (Figure 2). Correlations of assessment factors with overall course factor 2 indicate that both concept mapping and essay exams are related to their impressions of course fairness. Concept mapping and essay exams also appear to contribute to student beliefs that course work load is too much and too challenging as reflected in correlations with overall course factor 3 (Figure 2).

4.6 Sequential Relationships Among Assessment Factors

The ability to communicate understanding of a concept may be the ultimate test of that understanding. Hence, it was deemed important to assess the contributions of online quizzing and concept mapping to student perceptions of essay exams. To this end, online quiz and concept map factors were correlated with the essay exam factors (Figure 3). Similarly, it was deemed important to explore student impressions of the contribution of the online quizzing (which largely tests vocabulary) to the concept mapping factors. Student impressions that concept mapping is reasonable (factor 3) and that online quizzing is of general instructional value (factor 1) were positively correlated with their belief that essay exams are fair and that they were well prepared (factor 1). That concept mapping is reasonable (factor 3) was also correlated with the belief that essay exams improve understanding (factor 2). Finally, the desire for more online quizzing was correlated with the belief that essay exams were not challenging and that they were well prepared (Figure 3).

There were no significant correlations of the 3 possible online quiz factors with concept mapping factor 1 (maps goal oriented, required thought, promoted understanding), but there were significant correlations with concept mapping factors 2 and 3 (Figure 3). There was a significant positive correlation between online quiz factor 3, which generally expressed positive attributes of these quizzes, and concept mapping factor 2, which was positively correlated with the negative impression that mapping was too time-consuming and too frequent. This suggests that the students who like the online quizzes the most also had the most trouble with concept maps. There was also a positive correlation of online quiz factor 1 with essay exam factor 3. This suggests some direct relationship of the belief in online quizzing's general instructional value with some positive attributes of concept mapping.

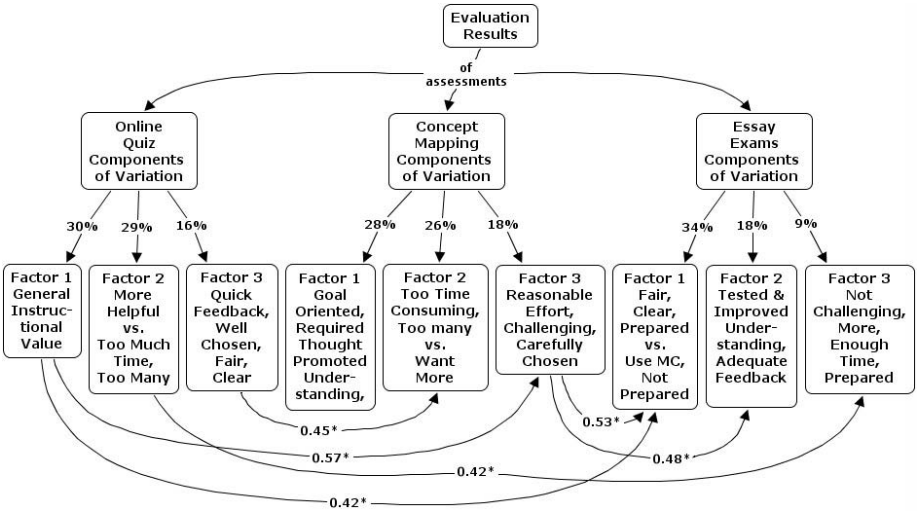

Figure 3. Concept map of evaluation results showing correlations of online quiz and concept mapping factors with essay exams and of online quiz factors with concept mapping factors. Percentages indicate variance explained by each factor. Correlation coefficients are shown for significant correlations between factors. Asterisks indicate significance of correlations: * = 0.05, ** = 0.01 and *** = 0.001. Six significant correlations out of 27 possible correlations are greater than expected by chance alone.

5 Hypotheses and Discussion

The results presented here examine some dimensions of variation in student opinions about a suite of three sequentially dependent assessment types and the overall course. For the online quizzing, the components of variation may be simplified to instructional value, helpfulness, and fairness. For the concept mapping, the major dimensions of variation can be summarized as promoting understanding, too time-consuming, and reasonableness. Essay exam factors reflected opinions related to bipolar fairness/preparedness, understanding, and not challenging. Finally, the overall course components can be summarized as reflecting course goals, fairness, and workload. Admittedly, it is risky to draw conclusions from a single small class of 23 students, so my conclusions are best stated as hypotheses for further study. Nevertheless, implicit in the design of this study were the a priori hypotheses that my alternative assessment would promote our overall course goals and that the sequential suite of assessments would culminate in a demonstration of deep learning as expressed by success on my essay exams.

5.1 Assessments and Deep Learning

Explicit goals of our course are to promote understanding of modern biology and science process, to help students develop their scientific reasoning skills, and to affect their worldview to include the interrelatedness of all life. Correlations of assessment factors with overall course factor 1 (Figure 2) suggests that all of the assessments contributed to the student's ability to communicate, understanding of biology, and broadened worldview.

Hypothesis 1: sequentially dependent assessments promote deep learning in science.

Research on student approaches to learning may help us explain how assessments can lead to conceptual understanding and metacognition. Student approaches to learning can be classified as surface and deep (Entwistle, 2001). Deep learners: 1) relate ideas to previous knowledge and experience; 2) look for patterns and underlying principles; 3) check evidence and relate it to conclusions; 4) examine logic and argument cautiously and critically; 5) are aware of the understanding that develops while learning; and 6) become actively interested in course content. Whereas, surface learners: 1) treat the course as unrelated bits of knowledge; 2) memorize facts and carry out procedures routinely; 3) find difficulty in making sense of new ideas presented; 4) see little value or meaning in either courses or tasks; 5) study without reflecting on either the purpose or strategy; and 6) feel undue pressure and worry about work (Entwistle 2001).

Clearly, we want our students to be deep learners and not surface learners. Surface learners will rarely develop deep conceptual understanding of any subject area, and they can be expected to do poorly on alternative assessments. Research on student approaches to learning has drawn attention to the significant influence of assessment procedures on learning and also led to identification of a third approach to learning often adopted by both surface and deep learner − the strategic approach. Strategic learners: 1) put consistent effort into studying; 2) manage time and effort effectively; 3) find the right conditions and materials for studying; 4) monitor the effectiveness of ways of studying; 5) are alert to assessment requirements and criteria; and 6) gear work to the perceived preferences of the teacher (Entwistle 2001). Thus, it appears that many students are keenly aware of the requirements for academic success and will modify their approaches to learning to ensure that they achieve that academic success, which is most often reflected by a high grade. Just as setting expectations can influence classroom behavior (Sufka and George, 2000), setting expectations about learning can influence student approaches to learning. If students accept expectations for learning and feel supported in achieving those expectations, strategic thinking may promote adoption of deep approaches to learning.

5.2 Alternative Assessment and Student Approaches to Learning

Correlations of concept mapping factors and essay exam factors with the fairness component of the overall course (factor 2; Figure 2) and with the workload component of the overall course (factor 3) suggest that student perceptions of these assessments have a great potential to affect student perceptions of the overall course. Student perceptions of unfairness and excessive workload might undermine student motivation and effort to perform the work of learning and undermine course goals.

Hypothesis 2: adequate support and acceptance of concept mapping and essay exams affect student approaches to learning and ultimately progress toward course goals.

Student acceptance of alternative assessments is key for achieving deep conceptual learning. Concept mapping and essay exams are relatively rare assessment types especially in science. Students may be conditioned to expect multiple choice exams and passive learning. Furthermore, students are spending fewer hours per week studying (Sufka and George, 2000), whereas learning in any domain requires a significant time commitment. Concept mapping and essay exams may require extra time because they are less familiar to students. Therefore, it is critical that students have adequate support in developing concept mapping and writing skills. Otherwise, they will consider them unfair and too demanding resulting in frustration and ultimately avoidance of learning. Student resistance can be overcome by providing adequate support and careful modeling of the concept mapping and writing processes (Quinn et al., 2004).

5.3 Concept Mapping for Concept Construction and Communication

Student perceptions that concept mapping was reasonable, challenging and carefully chosen (factor 3; Figure 3) was positively correlated to the fairness of essay exams and to the perception that they tested and improved understanding. This association is at least partly attributable to the fact that students were allowed to use their concept maps while writing their essays. However, it is also possible that the process of concept mapping helps students construct conceptual knowledge in preparation for good performance on the essay exams.

> Hypothesis 3: concept mapping is an effective tool for construction of conceptual knowledge and developing the ability to communicate deep understanding of that knowledge.

That concept mapping is an effective tool for deep learning has been well established (Novak, 2003). It appears that it not only promotes deep understanding, but also practice and mastery of the concept mapping process involve significant aspects of metacognition (Bransford et al. 2000). To an experienced grader, a student's concept map is a clear reflection of their understanding. Motivated students should also be able to detect and correct inconsistencies in their understanding as they map a concept. Such skills are fundamental to intellectual development and may help our students develop into reflective thinkers (Perry, 1968; King and Kitchener, 1994). Concept mapping is also an effective tool for developing effective written communication. Similar tools like mind mapping have long been used in composition classes for brainstorming prior to essay development. However, mind maps usually lack the hierarchical organization of concept maps and often converted to outlines prior to composition. Well constructed concept maps are as hierarchical as an outline, but retain crosslinks between hierarchies, which contribute to deep discussions of the concepts. Thus, concept maps can be effective tools for development of essays (Entwistle, 1995)

5.4 Surface Learning and Deep Learning

All online quiz factors were correlated with some essay exam factors, and online quiz factors 1 and 3 were correlated with concept mapping factors 3 and 2, respectively. Some of this relationship reflects the preference of some students for objective testing and their aversion to concept mapping, but the stronger correlation of quiz factor 1 with mapping factor 3 supports the value of online quizzing for concept mapping. Online quizzes were designed to promote surface learning.

> Hypothesis 4: assessments that promote surface learning such as online quizzes are a prerequisite for concept construction and communication.

Concept mapping may have the potential to be a bridge between surface learning and deep learning. However, it may be impossible to cross that bridge without first mastering the vocabulary for a concept. In other words, adequate surface learning may be a prerequisite for deep learning. Proponents of deep approaches to learning state that while the goal is deep understanding, it can not be achieved without first achieving a surface understanding (Entwhistle, 2001). The practical implication of this assertion is that students cannot explain or apply a concept without first learning the facts and vocabulary of a concept. So there must be some surface level expectation and assessment. Thus, assessments should be sequenced and coordinated to have maximum impact. They should first assure that students achieve surface learning including mastery of the basic vocabulary of the topic area. Then, assessments should promote construction of concepts including all of the major propositions and their interconnections. Finally, assessments should require construction of explanations of major concepts in preparation for communication about those concepts inside and outside the classroom. It is this last stage that should promote retention and transfer of conceptual knowledge about science that is the basis for scientific literacy (Halpern and Hakel, 2002). How faculty approach teaching affects curriculum design, teaching styles, and assessment choices. Many science faculty heeding the calls for science education reform have modified curricula and teaching to include more conceptual emphasis (Allen and Tanner, 2003; Udovic, 2002). However,

most of us have neglected to modify our assessment approaches and continue to rely on surface level assessments. Class sizes and time demands of alternative assessments can be impediments to assessment reform. However, we must find a way to reform our assessments if we hope to achieve our goals for scientific literacy.

6 References

Bransford, J.D., A.L. Brown, and R.R. Cocking (Editors). 2000. *How people learn: brain, mind, experience and school.* National Academy Press, Washington, D.C.

Educational Policies Commission, 1966. *Education and the Spirit of Science.* Washington, DC: National Education Association of the United States.

Entwistle, N. M. 1995. Frameworks for understanding as experience in essay writing and in preparing for examinations. *Educational Psychologist* 30:47-54.

Entwistle, N.J. 2001. Promoting deep learning through teaching and assessment. In Suskie, Linda, Editor. 2001. *Assessment to promote deep learning: insight from AAHE's 2000 and 1999 Assessment Conferences.* AAHE, Washington, D.C.

Halpern, D.F., and M.D. Hakel. 2002. Learning htat last a lifetime: teaching for long-term retention and transfer. In, D.F. Halpern and M.D. Hakel, Eds. *Applying the Science of Learning to University Teaching and Beyond*. New Directions for Teaching and Learning Number 89, Jossey-Bass, San Francisco.

King, P.M. and K.S. Kitchener. 1994. *Developing reflective judgment: understanding and promoting intellectual growth and critical thinking in adolescents and adults.* Jossey-Bass, San Francisco.

Lawson, A.E., 1988. A Better Way to Teach Biology. *The American Biology Teacher* Vol. 50:266-278.

Moore, R. 1996. Hands off Science. *American Biology Teacher* 58:387.

Novak, J.D. 2003. The promise of new ideas and new technology for improving teaching and learning. *Cell Biology Education* 2:122-132.

Perry Jr., W. G. 1968. *Forms of intellectual and ethical development in the college years: a scheme.* Holt, Rinehart & Winston, New York.

Quinn, J.J., J.J. Mintzes, and R.A. Laws. 2004. Successive Concept Mapping, Assessing Understanding in College Science Classes. *J. College Science Teaching* 33(3): 12-16.

Shamos, M.H. 1995. *The Myth of Scientific Literacy*. Rutgers University Press, New Brunswick, NJ.

Sufka, K.J., and M.D. George. 2000. Setting Clear and Mutual Expectations Affecting Learning Environments. *Liberal Education*, Winter: 48-52.