Proceedings of the Redesigning Pedagogy: Culture, Knowledge and Understanding Conference, Singapore, May 2007

USING CONCEPT MAP ASSESSMENT TO COMPLEMENT TRADITIONAL ASSESSMENT IN THE PHYSICS CLASSROOM

Theresa Thor

Raffles Institution, Singapore

Boo Hong Kwen

National Institute of Education, Nanyang Technological University, Singapore

Yap Kueh Chin

National Institute of Education, Nanyang Technological University, Singapore

ABSTRACT

The issues related to adopting concept map assessment to complement traditional assessment in a secondary 3 physics classroom in Singapore were examined. The 65 participants in the study were all high ability students. Training in concept mapping techniques was integrated into classroom lessons. Students sat for a traditional test before the concept map test on the same lesson units. The construct-a-map task was given with a concept list. The test maps were scored using a scoring method based mainly on the quality of propositions. The results suggest that concept map assessment may be used to complement traditional assessment in

the physics classroom. The measured achievement and analysis of the concept maps suggested that concept map assessment is a valid measure of student learning. Correlations of map scores with traditional test scores suggest that concept maps and traditional tests do not measure exactly the same aspect of knowledge and may have benefits of directing students to learn more meaningfully. The researcher posits that the construct-a-map task presented the high ability learners with the appropriate complexity and challenge. The training procedure, though effective, may not have been adequate to give all the students sufficient comfort in concept mapping.

INTRODUCTION

In Singapore schools, the shift has been towards using different modes of assessment to evaluate different dimensions of student learning. Despite this shift, the pen-and-paper assessments in physics continue to consist predominantly of traditional multiple-choice and short-answer tests. These assessments measure achievement with no consideration of the structure of a student's knowledge.

According to Novak (1977), the most important factor in cognitive learning is 'the extent to which meaningful learning has occurred and well differentiated, hierarchical cognitive structure has developed'. Based on Novak's theory, there should be a test instrument which teachers are able to use to assess the conceptual and propositional frameworks held by students.

PURPOSE OF STUDY AND RESEARCH QUESTIONS

Ruiz-Primo and Shavelson (1996) pointed out that reliability and validity information about different mapping techniques should be supplied before concept maps are used for assessment. The appropriateness of the cognitive demands of the concept-mapping task and the appropriateness of the scoring system used to evaluate the accuracy of the knowledge representation need to be considered (Ruiz-Primo, Schultz, and Shavelson, 1996). Research is also needed on students' attitude towards using concept maps and on the training techniques. The purpose of the study, therefore, was to investigate issues relevant to the adoption of concept map assessment to complement traditional multiple-choice and short-answer tests in the physics classroom. It focuses on the appropriateness of the construct-a-map task and a proposed scoring system based on the quality of propositions and map structure. The following research questions were addressed: (1) Did students in two different ability groups demonstrate different achievement levels in a common concept map test and traditional test? (2) What is the correlation between the students' concept map scores and traditional test scores? (3) Is there a difference in the concept maps of high and low achievers? (4) How did the students feel about concept mapping? (5) How effective is the training procedure adopted?

METHODOLOGY

Participants

The study involved two classes of secondary 3 (Grade 9) physics students in an all-boys secondary school in Singapore. All the students were high ability learners who had met stringent requirements to be admitted into the school. One class, referred to as group A in the study, comprised 36 students and was an advanced class for physics. Another class, referred to as group B, had 29 students and followed a regular course in physics. Both classes were taught by the same physics teacher who had 10 years of teaching experience.

Instrumentation

Defining the Mapping Technique

The students in the study were required to construct a map using a given concept list. As the construct-a-map technique elicits higher order cognitive processes (Ruiz-Primo et al., 2001), the researcher posited that it would impose suitable cognitive demand on the students. A completely open task where students construct a map using self-generated concepts was less desirable due to anticipated problems with scoring and comparability.

Defining the Response Format

The students constructed their maps on paper as this was most practical in a classroom situation. Observations from the training phase suggested the benefits of providing prompts in the test paper to remind students to plan, and work and rework their concept maps.

Defining the Scoring System

The scoring system was a modified relational scoring system from a technique developed by McClure and Bell (1990). Individual maps were scored by evaluating the separate propositions identified on the map. Each proposition was scored from zero to two depending on the quality of the proposition. A proposition inventory was developed and used as a guide during scoring. The inventory compiled the propositions provided by the students' maps and participant teacher's map. The scoring system also considered the organization of concepts in terms of the presence of higher-level structures within the concept maps. Zero to two points were awarded based on: (a) whether the superordinate/subordinate concepts in the students' maps make sense, and (b) the number of hierarchical levels identified on the maps. The scoring system is given in Table 1. The final raw score for each map was found by summing the scores of all the separate propositions and the structural organization score.

Component	Description	Point/s
Propositions	▪ Valid and complete key/mandatory proposition	2
	▪ Valid and complete 'other' proposition	1
	▪ Valid but incomplete proposition or proposition shows little or no understanding between concepts	0
	▪ Trivial or invalid/incorrect proposition	0
Hierarchy	▪ Does the map show hierarchy? Each subordinate concept is more specific and less general than the concept drawn above it.	1
	▪ At least 3 meaningful hierarchical levels	1
Others	▪ Missing directional arrows or missing enclosures for concepts (maximum 1 point penalty, deductible from total raw score)	- 1
Remarks ▪ no penalty if fewer than 20 of the given concepts are used ▪ total map raw score = total score for propositions + score for hierarchical organization		

Table 1. Scoring system

A conversion scheme was devised that converted the raw scores into scores on a scale point of 1-10. The conversion scheme is given in Table 2 below.

Raw scores of students	Converted scores of students (maximum: 10 points)
<7	1
7-8	2
9-10	3
11-12	4
13-14	5
15-16	6

17-19	7
20-22	8
23-25	9
>25	10

Table 2. Conversion scheme for raw scores

Constructing the Concept List

The researcher compiled a list of 20 key concepts in three ways: (a) asking the participant teacher to provide the concepts he thought were most important in the lesson units, (b) asking a physics education expert to provide the concepts he thought were most important in the lesson units, and (c) reviewing the concept lists and textbook material with the participant teacher.

The Traditional Test

The items in the traditional test paper were analyzed to determine the concepts and specifics required to answer them. There was overlap between the knowledge and understanding required in the traditional test, and the expected propositions in the concept map test except in instances where interpretation of diagrammatic representations was required. The traditional test comprised quantitative questions (33.3%), questions requiring qualitative reasoning (16.7%), questions on interpretations of diagrammatic representations (29.2%), questions requiring verbal explanations (12.5%), and questions requiring simple recall (8.3%).

The Questionnaire Survey

A simple questionnaire survey was designed and administered to assess the students' perception of concept mapping. The questionnaire used was similar in form to the questionnaire used in a study by McClure, Sonak, and Suen (1999) on concept map assessment of classroom learning. The seven items in the questionnaire are given in Table 3.

For each of the items below, indicate whether or not the statement is applicable to you by circling either ‘Yes’ or ‘No’.

1.	Drawing concept maps in physics class helped me learn better.	Yes	No
2.	Drawing concept maps in physics class was hard.	Yes	No
3.	Drawing concept maps in physics class was easy.	Yes	No
4.	Drawing concept maps in physics class was boring.	Yes	No
5.	Drawing concept maps in physics class was fun.	Yes	No
6.	I hate drawing concept maps in physics class.	Yes	No
7.	Any other comment or feedback.		

____
____
____

Table 3. Questionnaire survey

Logistics

Study procedures were carried out as part of the regular instructional programme in physics. The researcher adhered to the school assessment schedule and scheme of work. As such there were acknowledged constraints in the study such as those imposed by the assessment schedule and assessment plan, lesson sequence and timeframe, and even the format of the traditional test paper. The chronology of instruction and assessment can be found in Figure 1.

<table>
<tr><td colspan="7">Lessons continued according to scheme of work in the subject:
Physical quantities & measurement techniques/Light I/Light II/Light III/Waves/EM Spectrum/Sound</td></tr>
<tr><td>Term 1
Week 3</td><td>Term 1
Week 4</td><td colspan="2">Term 1
Week 8</td><td>Term 1
Week 10</td><td>Term 2
Week 1-2</td><td>Term 2
Week 6</td></tr>
<tr><td>Teacher Training</td><td>Student Training Phase 1</td><td colspan="2">Student Training Phase 2</td><td colspan="2">Student Training Phase 3</td><td>Data collection</td></tr>
<tr><td></td><td>Introductory student training with group work</td><td>Review of concept mapping skills; Timed-practice 1</td><td>Feedback on timed-practice 1</td><td>Timed-practice 2</td><td>Feedback on timed-practice 2</td><td>Review of concept mapping skills; Concept map test</td></tr>
<tr><td></td><td>Observation of students at work</td><td colspan="2">Observation of students at work, scoring concept maps & checking inter-rater consistency</td><td colspan="2">Observation of students at work, scoring concept maps & checking inter-rater consistency
Evaluation of student training</td><td>Survey</td></tr>
<tr><td colspan="7">Exploring and refining concept map task and scoring system</td></tr>
</table>

Figure 1. Chronology of concept map instruction and assessment

Training of Students

Procedures during Concept Map Training

According to Anderson and Huang (1989), students' learning of mapping skills is facilitated by using concept maps as part of instruction. Teaching students to use concept maps over several weeks also allowed the teacher to reflect upon the progress and to adjust instruction accordingly. The training procedure was thus weaved into the physics lessons. It comprised one group work session and two timed-practice sessions. The teacher followed detailed lesson plans during the entire course of the study to ensure consistency in instruction across the two classes.

Observations from Student Training

Observations of the students at work and of their concept maps were used to improve the concept mapping skills of the students. Feedback was given to the students on their concept mapping effort each time. Besides the technique of concept mapping, attention was drawn to the organization of concepts within a concept map and the quality of propositions. Examples of good concept maps were shown to the students to illustrate how much knowledge can be organized and represented in a map. To draw attention to the quality of propositions, the participant teacher explained the commonly rejected student propositions as well as student misconceptions on the lesson units.

DATA COLLECTION AND ANALYSIS

The researcher and participant teacher scored the students' maps separately and compared the scores. The inter-rater scoring was highly consistent.

The students' concept map and traditional test scores were collected. In addition the maps of 10 high achieving students and 10 low achieving students were identified for closer examination and analysis. The high and low achieving students were identified by first collapsing the groups A and B, and then ranking the students according to their overall physics scores in the last school semester. The 10 high achieving students were the top 10 scorers while the 10 low achieving students were the 10 lowest scorers. The participant teacher confirmed that the list of students was consistent with his observations of them in the classroom.

The assessment data in the last school semester revealed that group A consistently turned in stronger performances in traditional assessments. However group A was better than group B by a smaller margin in alternative assessments such as case studies and performance tasks. As such, analysis was also done by combining the data from the two groups of students. The combined group was referred to as group AB.

Research Question 1 concerned the achievement of two groups of students as measured by a concept map test and a traditional multiple-choice and short-answer test, while Research Question 2 concerned the correlation between the students' concept map and traditional test scores. To compare students' test scores first across assessment types, the descriptive statistics (mean, standard deviation) for the two types of assessments were computed. The correlations between the scores from the two types of assessments were determined. The pattern of correlations for the different groupings of students - AB, A, B - was examined as well.

Research Question 3 was addressed by analyzing the concept maps of 10 high and 10 low achievers. The concept maps were each analyzed for: (a) total number of propositions generated, (b) number of accurate propositions, (c) number of mandatory propositions, and (d) evidence of 'general-to-specific' pattern. The means for (a)-(c) and the percentage of students with the 'general-to-specific' pattern were computed for each group of achievers.

Research Question 4 on how the students felt about concept mapping was addressed by analyzing the students' responses to the questionnaire survey. The responses from both groups A and B were reported together. Recurring themes were identified from the data gathered from item 7 of the questionnaire. The data was classified into themes that emerged. This process was checked by the participant teacher to ensure that no division of the themes or new theme was needed.

To address Research Question 5, a simple evaluation of the student training was done. All the 65 individually constructed maps from groups A and B for timed-practice 2 were analyzed for: (a) percentage of students who used all the concepts provided, (b) percentage of students who used labeled links, and (c) number of valid propositions in the maps. The criteria set to deem the training programme as successful in training the students to construct concept maps was 80% for (a) and (b), and a minimum of 5 valid propositions in each map.

FINDINGS

Scores for Concept Map and Traditional Tests

Table 4 presents the mean scores and standard deviations for the two types of assessments taken by groups A and B. The lower mean scores for the concept map test for both groups suggest that the concept map test presented greater complexity and challenge to the students. Student responses received through the questionnaire survey showed that 61.3% of the students felt that drawing concept maps was hard. Only 22.6% of the students found it easy.

It was expected that students with greater competency in the subject matter would likely generate more propositions and be able to organize their concepts better than students with less competency. The results were consistent with expectations: Group A had higher mean scores than group B for both the concept map test and traditional test.

The standard deviations for both groups in the traditional test were similar. Both groups also had higher standard deviations for the concept map test compared with the traditional test. The concept map scores for group B were in fact more dispersed at a standard deviation of 24.07 compared with group A's standard deviation of 16.90. The greater dispersion in the concept map test scores for both groups is unlikely to be due to the scoring technique used for the concept map test as there was a high level of inter-rater consistency between the researcher and the participant teacher. The students' relative familiarity with traditional test questions and the drill-and-practice that students typically do to prepare for traditional tests may be one reason for the smaller standard deviations in the traditional test scores of both groups. A study by Bowden et al. (1992) demonstrated that with practice drills 'it becomes difficult to differentiate students on the basis of their traditional test scores' while Lin (1982) suggested that 'success in physics depends more on students' test-taking abilities rather than mastery of concepts'. The complexity of the concept map test together with the absence of drill-and-practice materials might have presented new challenges and higher cognitive demands on the students. The open-ended nature of the concept map test where students were required to articulate their understanding of the relationships between the given concepts and organize the map structure to show the superordinate-subordinate relationships seems to be a much more complex test for the students.

	Concept Map Test (Max = 100)[a]		Traditional Test (Max = 100)[b]	
Group	Mean	SD	Mean	SD
A (N=36)	60.00	(16.90)	79.67	(14.84)
B (N=29)	56.90	(24.07)	73.62	(14.72)

[a] Maximum score was calculated from concept map raw scores based on scale of 1-10 points and then converted to percentage.

[b] Maximum score was calculated based on total marking points of 24 and then converted to percentage.

Table 4. Means and standard deviations by group

The correlation between the concept map and traditional test scores was expected to be positive and weak-to-moderate in magnitude as the two tests do not measure exactly the same aspects of knowledge and skills. For example, the traditional test required strategies for problem solving and interpretation of diagrammatic representations. In addition, it was also possible for students to acquire practised efficiency in answering the traditional test questions from the large number of past examination and test questions that were available to the students in the school. Some students had asked the teacher how they could prepare for the concept map test.

Table 5 gives the correlation coefficients between the concept map and traditional test scores for the different groups. The correlation coefficient of combined group AB was moderate at .340. While this finding suggests that concept map and traditional tests measure 'overlapping and yet somewhat different aspects of declarative knowledge' (Ruiz-Primo, Shultz, and

Shavelson, 1997), it is interesting to note that group B on its own had a much higher correlation coefficient of .682 compared with the almost non-existent correlations of -.052 for group A. While students who did well in the traditional test can be expected to know some subject matter and concepts to make a good attempt in the concept map test, it is also possible that students may approach the concept map task differently. The students were placed in group A and group B at the start of the year based on their general academic performance across all subjects besides their specific performance in science. The students in group A may have a wider range of intelligences and higher order cognitive skills to deal with the complexity encountered in the concept map task. On the other hand, the students in group B may be doing little more than transferring linearly what they had applied in the traditional test into the concept map task. These are however conjectures that require further investigation.

Group		Correlation Coefficients between Concept Map and Traditional Test Scores
AB	(N = 65)	.340**
A	(N = 36)	-.052
B	(N = 29)	.682**

Note: The correlations are statistically significant as indicated by ** $p < .01$ (two tails).

Table 5. Pearson correlations compared by group

Examining the Maps of High and Low Achievers

The high achievers presented a greater amount of knowledge on their concept maps compared with the low achievers. They generated not only more propositions but also had more mandatory propositions and more accurate propositions. The maps reflected stronger content and organization of knowledge. The best maps showed more elaboration in the propositions, and creativity in organizing the structure of the map. Interestingly however, the high achievers recorded the same average number of trivial or inaccurate propositions as the low achievers. The students might be generating whatever they could think of within the timeframe given to them in the test. This may not be entirely surprising in a test situation, given the achievement-oriented nature of the students in the researcher's school. Figure 2 gives a concept map produced by one of the high achieving students. Table 6 gives a comparison of the data from the analysis of the concept maps of the high achievers and low achievers.

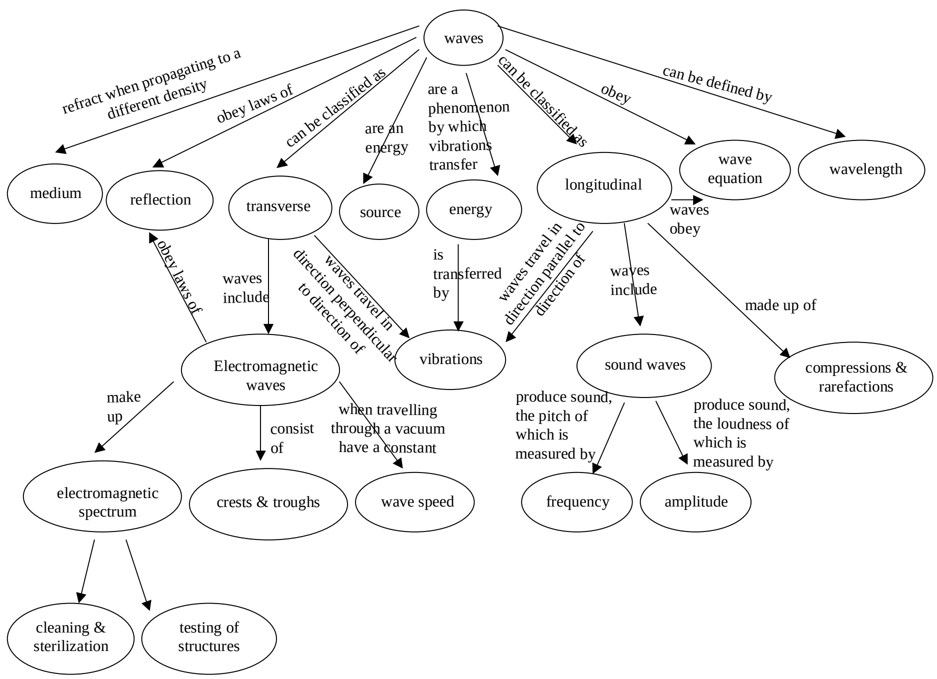

Figure 2. Concept map of a high achiever

	High achievers	Low achievers
Group mean of total propositions generated	24	17
Group mean of mandatory propositions	4	1
Group mean of accurate propositions	16	9
Group mean of % accurate propositions	68.2	55.3
Group mean of trivial or inaccurate propositions	8	8

% maps with general-to-specific pattern	90	20

Table 6. Data from concept maps of high and low achievers

The low achievers not only generated fewer propositions, they also had fewer mandatory propositions and a lower percentage of accurate propositions. Their concept maps generally showed a high proportion of basic factual statements that were incomplete, content-inaccurate or not meaningful. There was also little organization with an evident lack of the general-to-specific pattern in the map structure of 8 out of the 10 students. Figure 3 shows the concept map produced by one of the low achieving students.

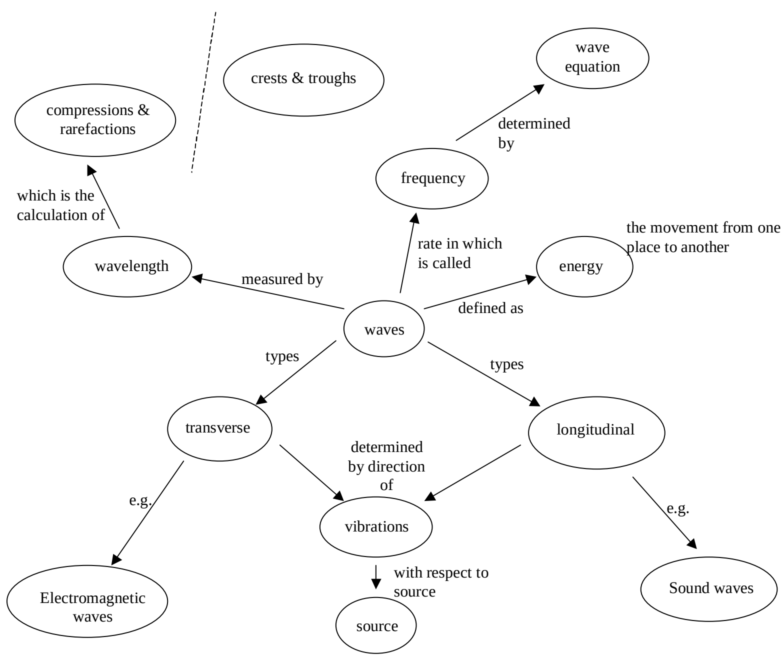

Figure 3. Concept map of a low achiever

Student Perception of Concept Mapping

In general, more students had negative perceptions of concept mapping. 63.5% of the students did not agree that concept mapping helped them learn better. 61.3% found concept mapping hard. Only 22.6% found it easy. 33.9% found concept mapping fun while 66.1% hated it.

The students' feedback from the survey suggested that they might not be entirely comfortable with concept mapping and might not all see the significance of concept mapping as a useful tool. It is possible that some may be looking for the kind of fluency that they had with traditional test items. The relative newness of the assessment tool, and the absence of practice drills that have typically given many of these students the ease in handling traditional test items in the Singapore context, may have caused the increased student anxiety that the researcher and participant teacher detected. This, together with the training approach that was adopted, which comprised two timed-practice sessions after which students' concept maps were scored, could have contributed to the largely negative student perception of concept mapping. It is not possible to say in the study if the negative perception of concept mapping impacted the students' performance in the concept map test.

Training Procedure

Results from the analysis of all 65 concept maps from timed-practice 2 indicated that (a) 100% of the students in group A and 93.1% of the students in group B used all the concepts provided, (b) 100% of the students in both groups used labeled links, and (c) 86.1% in group A and 93.1% in group B had at least 5 valid propositions in their maps. The training procedure was thus technically successful in training the students to construct concept maps. The training procedure, which incorporated timed-based practices, allowed the students to get a sense of working within a given timeframe. In addition, the participant teacher grew more confident in scoring concept maps and was able to suggest improvements.

CONCLUSIONS AND RECOMMENDATIONS

The study explored the potential of using concept map assessment to complement traditional assessment in the physics classroom by investigating issues relevant to its adoption. The findings suggest that concept map assessment can be practically applied in the physics classroom to complement traditional assessment with some caution. The measured achievements as well as the characteristics of the concept maps observed, provided evidence to support the validity of the construct-a-map task and scoring system used. Inter-rater consistency, although not a key focus in the study, was not an issue with the use of a proposition inventory; the time-cost was not unlike that in the scoring of a traditional test. The construct-a-map task seems to present suitable challenge for the high ability students in the school. The open-ended nature of the construct-a-map task that requires students to organize the concepts and articulate their understandings in their own words may move the students away from the drill-and-practice way of learning to more meaningful learning. The students can be easily trained on the technical procedures of concept mapping. Incorporating training activities into the lessons allows students to gain increasing competence in concept mapping.

The findings suggest a general strategy for using concept map assessment to complement traditional assessment in the physics classroom. The recommendation is to use concept maps first to facilitate learning and to give lots of practice to build fluency before introducing concept map assessment. The teacher needs to not only teach the technical procedures of concept mapping but also understand better how the students approach the concept mapping task in order to build the skills required for successful concept mapping.

A construct-a-map task designed on any lesson unit directs the students to organize the key concepts and specifics, and to explicitly articulate the relationships between them. As well as promoting deep understanding and creativity, this assessment approach supports constructivist learning. Students learn to be more active participants in building their own scientific knowledge structure. Additionally, use of the construct-a-map task allows both gaps in knowledge and misconceptions to be accessed and addressed. This information on learning may be given to both students and teachers. As a formative assessment tool, the teachers may use concept maps to study alternative conceptions in the physics classroom and revise classroom practices while the students can use the feedback to monitor their own learning.

Concept map assessment measures the structure dimension of knowledge, which is not addressed in traditional multiple-choice and short-answer tests. In this way the concept map as an instrument to complement traditional multiple-choice and short-answer tests in physics is consistent with the broader notion of science achievement and is aligned to current thinking on meaningful learning. To quote Wandersee (1990), '… what is basic to making a concept map for a piece of scientific knowledge is the ability of the mapper to identify and relate its salient concepts to a general, superordinate concept and this requires an understanding of what constitutes a science concept'. Hence concept map assessment promotes higher order thinking skills and can impact science teaching and learning in a positive way.

REFERENCES

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). Urbana-Champaign: Center for the Studying of Reading, University of Illinois at Urbana-Champaign. (ERIC Document Reproduction Service No. ED 310 368).

Bowden, J., Dall'Alba, G., Martin, E., Laurillard, D., Marton, F., Masters, G., et al. (1992). Displacement, velocity, and frames of reference: Phenomenographic studies of students' understanding and some implications for teaching and assessment. *American Journal of Physics, 60*(3), 262-269.

Lin, H. (1982). Learning physics versus passing courses. *The Physics Teacher*, *20,* 115.

McClure, J. R., & Bell, P. E. (1990). *Effects of an environmental education-related STS approach instruction on cognitive structures of preservice science teachers.* University Park, PA: Pennsylvania State University. (ERIC Document Reproduction Service No. ED 341 582).

McClure, J. R., Sonak B., & Suen, H. K. (1999). Concept Map Assessment of Classroom Learning: Reliability, Validity, and Logistical Practicality. *Journal of Research in Science Teaching, 36*(4), 475-492.

Novak, J. D. (1977). *A theory of education*. New York: Cornell University Press.

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., Schultz, E. S., & Shavelson, J. R. (1996, April). *Concept map-based assessments in science: An exploratory study.* Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Ruiz-Primo, M. A., Schultz, E. S., & Shavelson, J. R. (1997, March). *On the validity of concept map-based assessment interpretation: An experiment testing the assumption of hierarchical concept maps in science*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Ruiz-Primo, M. A., Schutlz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching, 38*(2), 260-278.

Ruiz-Primo, M. A., Shavelson, R.J., Li, M., & Schultz, S.E., (2001). On the validity of cognitive interpretations of scores from alternative concept-mapping techniques. *Educational Assessment*, *7*(2), 99-141.

Wandersee, J. H. (1990). Concept mapping and the cartography of cognition. *Journal of Research in Science Teaching, 27*(10), 923-936.