Student-Centered Reliability, Concurrent Validity and Instructional Sensitivity in Scoring of Students' Concept Maps in a University Science Laboratory*

Osman Nafiz KAYA[1] & Ziya KILIÇ[2]

[1] Gazi University, Gazi Faculty of Education, Department of Science Education, Ankara, TURKEY.

[2] Gazi University, Gazi Faculty of Education, Department of Chemistry Education, Ankara, TURKEY.

Abstract: Student-centered approach of scoring the concept maps consisted of three elements namely symbol system, individual portfolio and scoring scheme. We scored student-constructed concept maps based on 5 concept map criteria: validity of concepts, adequacy of propositions, significance of cross-links, relevancy of examples, and interconnectedness. With respect to the concurrent validity of scoring the concept maps in this study, correlation coefficients were computed between total scores of students' concept map and their scores of Prior Knowledge Test (PKT), and Achievement Test (AT) involved in the concepts of science laboratory experiments. The mean scores of students' pre-lab concept maps correlated much better with their scores of PKT ($r= .615$, $p < 0.01$), and also mean scores of students' post-lab concept maps correlated with their scores of Achievement Test ($r= .478$, $p < 0.05$). In regard to instructional sensitivity, the statistical analysis based on the comparison of both total and interconnectedness scores of pre- and post-lab concept maps indicated that there were significant differences favoring the post-lab concept maps ($p < 0.01$).

Introduction

The concept map has known to be multifunctional in science instruction. Novak (2001) stated, "When concept maps are used in instruction, they can also be used for evaluation. There is nothing written in stone that says multiple choice tests must be used from graduate school through university, and perhaps in time even national achievement exams will utilize concept mapping as a powerful evaluation tool" (p. 9). Several instructors have

* Poster presented at 18th International Conference on Chemical Education "Chemistry Education for the Modern World", İstanbul, TURKEY, 2004.

used pre- and post-lecture concept maps with teacher-supplied concepts, at intervals ranging one period to four-six weeks, in units of science instruction such as physics (e.g., electricity), biology (e.g., photosynthesis, marine life), and chemistry (e.g., oxidation-reduction, atomic structure), for describing the development of students' conceptual understanding over a period of time (e.g., 3 weeks, two semesters) (Hegarty-Hazel, 1991 a, b; Markow & Lonning 1998; Martin, Mintzes, & Clavijo, 2000; Pearsall, Skipper, & Mintzes, 1997; Regis, Albertazzi, & Roletto, 1996; and Wallace & Mintzes, 1990). These authors have analyzed concept maps on several criterions, which may include some or all of the following: concepts, propositions, relationships, cross-links, hierarchy, general to specific, branching, integration, differentiation, and interconnectedness.

What impact student-constructed maps had on conceptual learning was the primary purpose of the foregoing studies. Concept map was used as a powerful tool to assess student conceptual understanding and to promote concept development. However, these studies did not directly measure conceptual understanding based on science concepts structured through concept maps. Other issues are: firstly, except for Peersall et al., (1997) and Martin et al., (2000) studies, all other conceptual understanding studies using pre- and post-concept maps employed teacher given concept labels, which in our view is not a valid approach for measuring conceptual understanding because this approach restricts the use of students' personal knowledge and understanding **(**Ruiz-Primo & Shavelson 1996). Secondly, the focus in most studies was whether or not concept mapping would bring about better achievement scores in traditional tests. Learning from these previous studies, our study (1) devoted a long period of training in concept mapping skills before research activities; (2) used students' personal knowledge (their own concept labels) for constructing the concept maps; (3) developed a symbol system together with students for scoring concept maps, which sought to address the issue of reliability (objectivity); (4) used student's individual portfolio for tracing the conceptual change; (5) included the missing criterion of "example of concepts"; (6) added the new category named "interconnectedness"; (7) used pre- and post-laboratory concept maps as a measure of conceptual learning instead of traditional tests; and (8) identified students' alternative conceptions in pre- and post-laboratory concept maps.

Purpose

Taking the points above into consideration, the aim of this paper is to investigate the concurrent validity and instructional sensitivity of students' concept maps as vehicles for exploring their conceptual understanding. *Concurrent validity* is the correlations between concept map scores and other measures of students' achievement (Ruiz-Primo & Shavelson, 1996). *Instructional sensitivity* is the comparisons of concept maps before and after instruction with respect to the number of components in concept maps (Ruiz-Primo & Shavelson, 1996).

Significance of this study

Until a couple years, although concept mapping has been widely accepted as a valid tool in assessing students' knowledge structure, studies have been focusing more on the effects of concept mapping on students' achievement in science. The distinct difference between this study and the earlier studies is the use of a student-centered approach as a reliable method of scoring concept map instead of the interrater reliability or inter-scorer agreement used in many previous studies (Nakhleh & Krajcik 1991; Novak & Musonda, 1990; Rice, Ryan & Samson, 1998; Ruiz-Primo, Schultz, Li & Shavelson 2001; Rye and Rubba, 2002).

Research Questions

1. With respect to the concurrent validity;

- Are there significant correlations between scores of students' pre-lab concept maps and their scores of prior knowledge test?
- Are there significant correlations between scores of student' post-lab concept maps and achievement tests in science laboratory topics?

2. Are there significant differences between the total and interconnectedness scores of students' pre- and post-lab concept maps with respect to the instructional sensitivity?

Methods

Participants

A total of 20 students (7 males and 13 females), ages 20-21, were randomly selected from five science laboratory classes taught in the second semester of the 2000-2001 academic year in the Faculty of Gazi Education, Ankara, Turkey.

Procedures and Laboratory Design

The science laboratory course was 8 weeks long. The class met once a week for 3-hours, which constituted 24 hours of laboratory instruction. Students were grouped into small groups of three consisting of both male and female to facilitate collaborative learning. Science laboratory learning was planned for prospective middle school science teachers. All of the students were given only the names of the general concepts related to lab experiments which would be performed, for example, chemical reaction, reaction heat, and acids-bases and neutralizations. There are 8 predetermined lab topics and students worked in groups of 2-3. Each group worked on a different topic every week on a predetermined rotational schedule. The laboratory topics were: (1) the energy of reactions (reaction heat), (2) acids-bases (buffer solutions), (3) solution chemistry (solubility), (4) basic cell structure and functions, (5) cellular reproduction (meiosis and mitosis), (6) photosynthesis, (7) resistance and (8) magnetism.

Training in Concept Mapping

The training sessions involved teaching students how to construct and score concept maps using several chemistry topics, which took a-4 hours of laboratory time. The first training session consisted of explaining and discussing what was meant by a concept and the associated terminologies such as concept label, linking words, hierarchy, examples of concept, proposition, and cross-link. In the second laboratory session, the instructor did a practice exercise with students on the particulate nature of matter by selecting the central or main concept from a list of concepts, displaying examples of hierarchical and non-hierarchical concept maps, and constructing a concept map. The third laboratory session was, therefore, devoted to assess the concept maps constructed by the students at home, to help students identify valid and invalid propositions, cross-links between concepts, and examples of concepts and also to practice the symbol system, individual portfolio and scoring schemes.

Scoring the Concept Maps

We examined student-constructed concept maps using 5 concept map criterions: validity of concepts, adequacy of propositions, significance of cross-links, relevancy of examples, and interconnectedness. The same students' pre- and post-lab concept maps are shown in Figures 1 and 2. The researchers provided each student a portfolio to keep his or her pre- and post-laboratory concept maps. On the first page of each portfolio, the instructor placed a sheet containing two tables: One for assessing students' pre-laboratory concept maps and the second for assessing students' post-laboratory concept maps. By comparing the scores tabulated for each pre- and post-laboratory concept maps, the instructor was able to assess each student's conceptual understanding. Moreover, the student was also able to monitor his/her conceptual growth, keep track of scores, and check their own concept map scores.

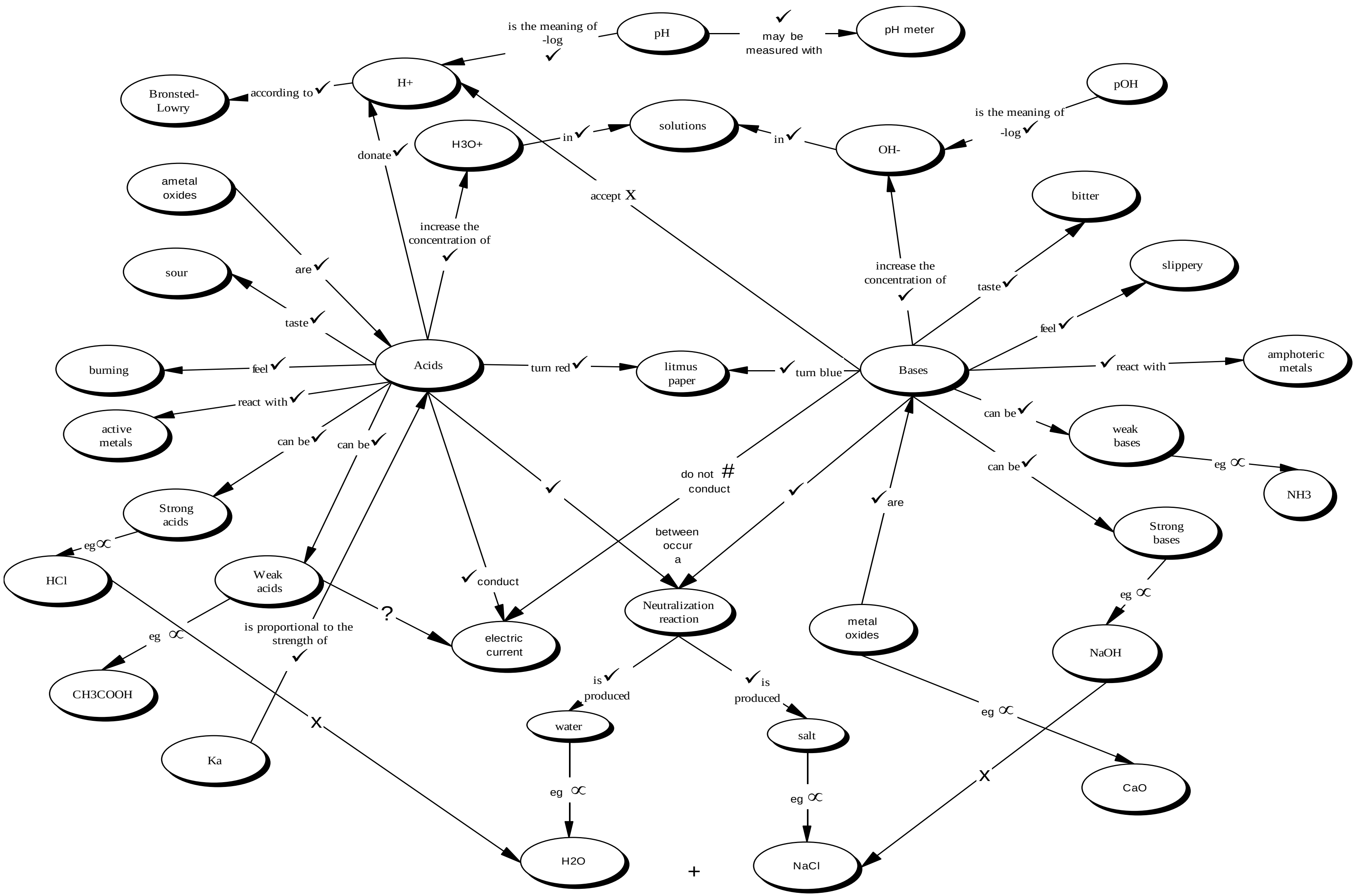

Figure-1. A student's pre-laboratory concept map for the experiment of acids-bases and neutralization

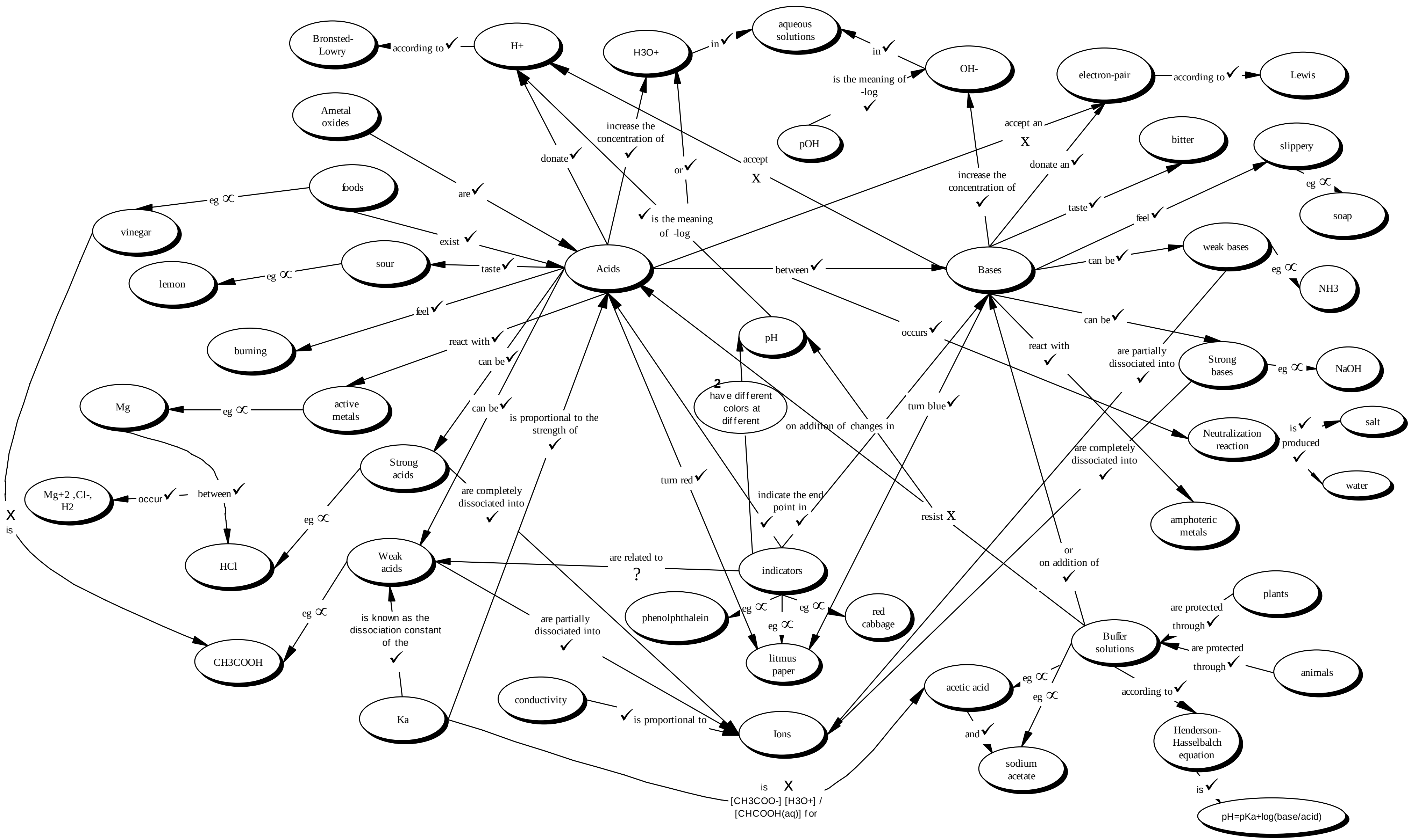

Figure-2. The same student's post-laboratory concept map for the experiment of acids-bases and neutralization

The symbol system for assessing student constructed pre- and post laboratory concept maps:
(√) a valid proposition
(**X**) a valid and significant cross-link
(α) a valid example of concept
(**#**) an alternative conception
(**?**) unclear proposition, unlabelled relationships between concepts, and incorrect isolated concepts and linking words

Scoring ways for pre- and post- laboratory concept map of a student are below.

**The total score* = [(valid propositions x 1 point) + (valid and significant cross-links x 10 points) + (valid examples x 1 point)]

**The value of interconnectedness* = [(valid and significant cross-links x 10 points) / (valid concepts x 1 point) x 100]

Instruments

Prior to the beginning of this study, Prior Knowledge Test (PKT), a-20 item multiple-choice test was developed by the researchers with the help of two experts in the lab topics selected for this study and administered to all students for determining students' prior knowledge related to the science topics. As a post-instructional test, Achievement Test of Science Laboratory Topics (AT), a-30 item multiple-choice test was developed by the researchers with the help of the same experts in the selected lab topics and administered to all students for determining students' achievements related to science laboratory experiments. Cronbach's alpha reliabilities of these tests were respectively 0.73 and 0.78.

Results

We calculated the total average score of each student' pre-lab concept maps scores (TASPreLCMs) for all lab topics. The descriptive statistics of the prior knowledge test and total average score of pre-lab concept maps are given in Table 1. It was computed correlation coefficients between PKT scores and TASPreLCMs. The total average scores of students' pre-lab concept maps correlated highly with their scores of Prior Knowledge Test ($r= .615$, $p < 0.01$). This statistical analysis is shown in Figure 3 as a bivariate scatterplot graphic.

Table 1. Means, Standard Deviations, Minimum and Maximum Values of the Prior Knowledge Test and Total Average Score of Pre-Lab Concept Maps.

	N	Min.	Max.	Mean	Std. Deviation
PKT	20	5.00	15.00	9.60	2.98
TASPreLCMs	20	53.00	81.00	69.55	7.79

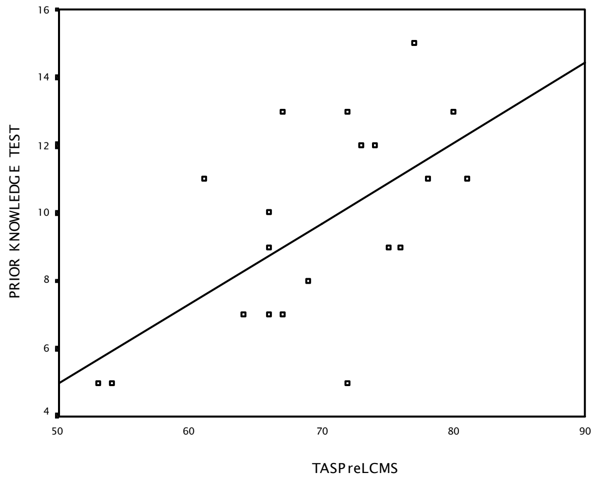

Figure-3. Bivariate scatterplot of continuous variables: TASPreLCMs against Prior Knowledge Test scores.

It was also calculated the total average score of each student' post-lab concept maps scores (TASPostLCMs) for all lab topics. The descriptive statistics of the Achievement Test (AT) and total average score of post-lab concept maps are given in Table 2. The correlation coefficient was computed between AT scores and TASPostLCMs. The total average scores of students' post-lab concept maps correlated with their scores of Achievement Test ($r = .478$, $p < 0.05$). This statistical analysis is shown in Figure 4 as a bivariate scatterplot graphic.

Table 2. Means, Standard Deviations, Minimum and Maximum Values of the AchievementTest and Total Average Score of Post-Lab Concept Maps.

	N	Min.	Max.	Mean	Std. Deviation
AT	20	11.00	29.00	21.45	5.09
TASPostLCMs	20	77.00	114.00	96.25	8.29

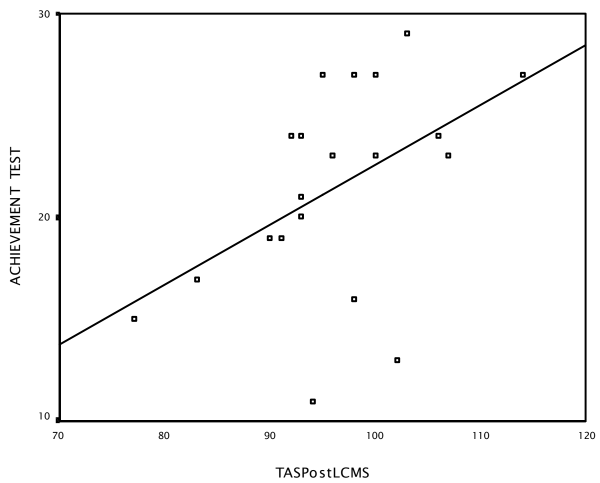

Figure-4. Bivariate scatterplot of continuous variables: TASPostLCMs against Achievement Test scores.

We also analyzed whether or not there are significant differences between the total mean scores of pre- and post-lab concept maps by using paired sample t-test with respect to instructional sensitivity. These analyses given in Table 3 indicated that there are significant differences favoring the students' post-lab concept maps ($p < 0.001$).

Table 3. The total mean scores and standard deviations (below) of pre- and post- laboratory. concept maps for the science laboratory experiments.

Pre- Lab. Concept Maps (n= 20)		Post- Lab. Concept Maps (n= 20)	*t*	*p*
1.	48.15 (9.62)	71.60 (16.22)	6.268	.000
2.	44.95 (10.57)	57.20 (14.25)	5.854	.000
3.	40.55 (15.18)	56.35 (16.67)	5.029	.000
4.	47.95 (13.04)	63.95 (11.01)	6.367	.000
5.	42.70 (11.06)	63.00 (14.67)	7.500	.000
6.	42.20 (6.91)	71.60 (9.37)	11.352	.000
7.	44.00 (10.33)	63.90 (15.20)	6.223	.000
8.	54.20 (13.48)	75.15 (12.92)	5.226	.000

Summary and Conclusions

The first aim of this study was to develop and examine a different way of obtaining reliability in scoring of the student-constructed concept maps instead of interrater reliability used in most previous studies. Hence, we used a student-centered reliability consisting of a symbol system, individual portfolios and scoring schemes. Based on our results, all three elements supported each other with respect to reliability in scoring the concept maps. We believe that if the interrater reliability is independently done based on the total scores of students' concept maps for more than one criterion by the raters, reliability in scoring the concept maps may not be reliable. The second purpose of this study was to examine the validity in scores of the concept maps in two different ways. We examined the correlation coefficients between the total scores of students’ concept maps and relevant test scores with respect to concurrent validity. These correlation coefficients were high especially for the scores of pre-lab concept maps and prior knowledge test. According to these correlation coefficients, we think that these different assessment tools may point out similar things with respect to the students' knowledge structure. Also, analyses with respect to the instructional sensitivity indicated that there was a statistical difference in growth in students' knowledge structure. It could be claimed that the reliability in scoring of students' concept maps certainly affects the validity in scoring of the concept maps.

References

Hegarty-Hazel, E., & Prosser, M. (1991a). Relationship between students' conceptual knowledge and study strategies. Part 1: Student learning in physics. International Journal of Science Education, 13(3), 303–312.

Hegarty-Hazel, E., & Prosser, M. (1991b). Relationship between students' conceptual knowledge and study strategies. Part 2: Student learning in biology. International Journal of Science Education, 13(4), 421–429.

Markow, P.G., & Lonning, R.A. (1998). Usefulness of concept maps in college chemistry laboratories: Students' perceptions and effects on achievement. Journal of Research in Science Teaching, 35, 1015-1029.

Martin, L.B., Mintzes, J.J., & Clavijo, E.H. (2000). Restructuring knowledge in biology: cognitive processes and metacognitive reflections. International Journal of Science Education, 22, 303-323.

Martin, L.B., Mintzes, J.J., & Clavijo, E.H. (2000). Restructuring knowledge in biology: cognitive processes and metacognitive reflections. International Journal of Science Education, 22, 303-323.

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

Novak, J.D. & Musonda, D. (1990). A twelve-year longitudinal study of science concept learning . American Educational Research Journal, 28, 117-153.

Pearsall, N.R., Skipper, J.E.J., & Mintzes, J.J. (1997). Knowledge restructuring in the life sciences: A longitudinal study of conceptual change in biology. Science Education, 81, 193-215.

Regis, A., Albertazzi, P.G., & Roletto, E. (1996). Concept maps in chemistry education. Journal of Chemical Education, 73, 1084-1088.

Rice, D., Ryan, J., & Samson, S. (1998). Using concept maps to assess student learning in the science classroom. Must different methods compete? Journal of Research in Science Teaching, 35, 1103-1127.

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

Ruiz-Primo, M.A., Schultz, 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, 260-278.

Rye, J. A., & Rubba, P. A. (2002). Scoring concept maps: An expert map-based scheme weighted for relationships. School Science & Mathematics, 102, 1, 33--45.

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