Concept Maps: Making Learning Meaningful
Proc. of Fourth Int. Conference on Concept Mapping
Viña del Mar, Chile, 2010

STRUCTURAL ANALYSIS OF CONCEPT MAPS TO EVALUATE THE STUDENTS' PROFICIENCY AS MAPPERS

Sérgio Noronha Silva Jr, Jerson Geraldo Romano Jr & Paulo Rogério Miranda Correia, Universidade de São Paulo, Brasil

Email: prmc@usp.br

Abstract. Mappers' proficiency on concept mapping is essential to achieve all promised rewards we want to find when using Cmaps to represent and share knowledge. Training sessions for beginners should be explored to overcome the naive use of this technique. This work proposes a quantitative structural analysis of the Cmap fine morphology to determine the skill level of students during the ACH 0011 Natural Science course. Sixty nine Cmaps were analyzed and the results are presented by using descriptive statistic and hierarchical cluster analysis (multivariate exploratory approach). We characterized 5 clusters containing Cmaps with distinguishable features, and the Cmaps made by beginners (class 1) were quite different from those prepared by these students in class 15. Among the 8 parameters used to describe the Cmap fine morphology, propositional density, initial concepts with multiple propositions, and final concepts with multiple propositions can be considered as potential "fingerprints" to monitor the students' expertise on concept mapping, since in our case, they increased when the students' experience with concept mapping increases.

1 Introduction

Concept maps (Cmaps) were proposed by Novak and colleagues during the early 1970s. They can be defined as a set of concepts embedded into a propositional framework. Concept mapping plays a key role as a tool to represent knowledge held by a learner, and also the structure of knowledge in any subject field (Novak, 2010).

In spite of being a well-established technique, widely used for educational and corporate purposes with a broad range of goals (Coffey et al, 2004; Coffey, 2006; Fourie & Westhuizen, 2008; Kyrö & Niskanen, 2008; Novak, 2010; Torres & Marriott, 2009), some obstacles still need to be overcome. The apparent ease of production of Cmaps is tempting to beginners and explains their popularity. However, naive use of concept mapping may cause few (or none!) of the expected benefits, and such experiences may be playful and funny at best (Correia, Infante-Malachias & Godoy, 2008). Some works in the literature show that many of the difficulties observed with the use of Cmaps derive at least in part from the inappropriate use of the technique, inadequate training for users and trainers, and a general failure to recognize the importance of the tool's theoretical foundations (Correia, Infante-Malachias & Godoy, 2008; Cañas & Novak, 2006).

Cmaps can be analyzed in different ways for different purposes. Figure 1 presents our understanding about three complimentary evaluations that can be useful to capture all information available in Cmaps. The concepts in Figure 1 are organized as follows: Cmap key parts (corner A), Cmap analyses using a semantic approach (corner B), Cmap structural analysis (corner C), and considerations about the mappers' proficiency (corner D).

This framework expresses our beliefs about the importance of training mappers to make good Cmaps. Training is essential to ensure the experience of the rewards of concept mapping, and purposeful activities should be designed for this goal.

The aim of this work is to examine the relationships between network morphology of Cmaps (Fig. 1, corner C) and the mappers' experience (Figure 1, corner D). We hypothesize that the advancement of mappers' expertise will increase the complexity of the Cmap structure. Therefore, an 8-paratemeter set was devised to develop a quantitative appreciation of the structural features of Cmaps.

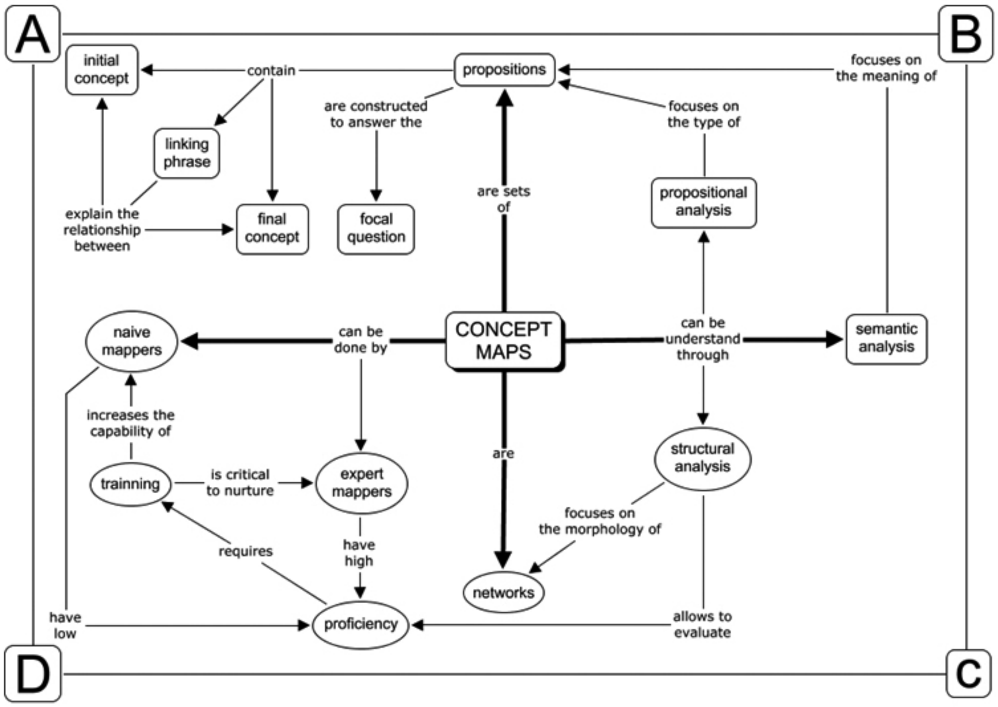

Figure 1. The relationship among the Cmap key parts (A), analyses using a semantic (B) and structural (C) approaches, and the mappers' proficiency (D). The core idea of this work involves the links between structural analysis (B) and mappers' proficiency (D)

2 Research procedures

2.1 Setting and data collection

The Cmaps (n=69) considered in this work were obtained during the ACH 0011 Natural Sciences course, which is offered for all first-year students at Escola de Artes, Ciências e Humanidades (School of Arts, Science and Humanities at São Paulo University). The main goal of this course is to provide a comprehensive view of the impact caused by scientific and technological development in our society (Correia et al, 2010; Santos, 2007). Sixty students from four different undergraduate courses were grouped in one classroom, for two-hour weekly classes over a period of fifteen weeks. Considering the course's introductory scope and the diverse audience, the challenge to meet students' expectations required innovative methodological strategies.

Students were introduced to concept mapping and trained to elaborate both manuscript and digital Cmaps using the Cmaptools software at the beginning of the ACH 0011 course. Half-structured Cmaps (HSCmaps), expanded collaborative learning (ECL) and propositional clarity table (PCT) are innovative approaches that enhanced the training session (classes 1-4), as described in the literature (Correia, Infante-Malachias, & Godoy, 2008). Two different sets of Cmaps were considered in this work: collaborative Cmaps (n=14) obtained during class 1 (naive mappers), and individual Cmaps (n=55) obtained during the class 15 (experienced mappers). Table 1 summarizes the most relevant information about the conditions used for asking these Cmaps. It should be highlighted that the subject in both cases (class 1 and 15) were familiar to the students, and that the main difference between them is the students' expertise on concept mapping.

	Class 1 Cmaps	Class 15 Cmaps
Number of collected Cmaps	14	55
Authorship	Collaborative, 2- or 3-student groups	Individual
Number of concepts	11-14	9
Focal question	*Why is so hard to get a place in the undergraduate courses offered by USP?*	*How does bioethics regulate the relationship between science and society?*
Students' familiarity about the subject	High	High
Students' proficiency on concept mapping	Naive	High

Table 1. Comparison about the Cmaps considered in this work

2.2 *Structural analysis of HSCmaps*

Among several works dealing with structural analysis of Cmaps (BouJaoude & Attieh, 2008; Gerstner & Bogner, 2009; Kinchin & Alias, 2005), there is a outstanding paper published 10 years ago (Kinchin, Hay & Adams, 2000). The authors proposed a qualitative approach to analyze the gross morphology of Cmaps. The main patterns (spoke, chain, and net structure) were related to the students' understanding about the mapped and their learning preferences. This qualitative approach inspired us to develop a quantitative description of the Cmap networks. We developed a set of 8 parameters that capture all structural features of Cmaps. These parameters are based on the fundamental building blocks J, K, and L shown in Figure 2.

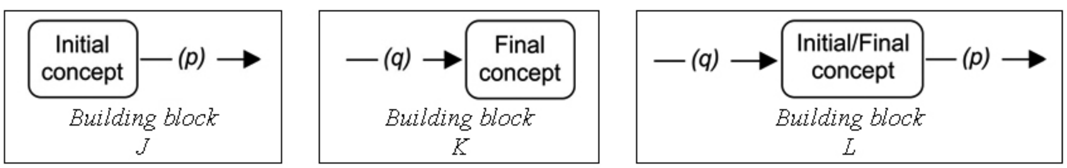

Figure 2. Fundamental building blocks (J, K, and L) of Cmaps' network considered for structural analysis

The parameters for structural analysis are described from the variables represented by J, K, L, p and q letters, which are found in the proposed fundamental building blocks (Figure 2). Table 2 presents the 8 parameters devised for the quantitative structural analysis, the notation adopted throughout this work, and the computation of each parameter.

- J: initial concept structure, characterized by a concept with arrow(s) pointing to another concept(s).
- K: final concept structure, characterized by a concept with arrow(s) pointing to it.
- L: initial and final concept structure, characterized by a concept with arrow(s) pointing to another concept(s) and arrow(s) pointing to it.
- p: number of propositions formed from an initial concept.
- q: number of propositions finishing in a final concept.

Parameter for structural analysis			Correspondence with structural building blocks	Computation
	Description	Notation		
	Concepts (boxes)	C	-	-
	Propositions (arrows)	P	-	-
1	Propositional density	PD	-	PD=C/P
2	Only initial concepts	OIC	J	OIC=J/C
3	Only final concepts	OFC	K	OFC=K/C
4	Initial/final concepts	IFC	L	IFC=L/C
5	Initial concepts	IC	J+L	IC=(J+L)/C
6	Initial concepts with multiple propositions	MIC	Jm+Lm=J+L, when (p≥2)	MIC=(Jm+Km)/C
7	Final concepts	FC	K+L	FC=(K+L)/C
8	Final concepts with multiple propositions	MFC	Km+Lm=K+L, when (q≥2)	MFC=(Km+Lm)/C

Table 2. Parameters for the quantitative structural analysis of Cmaps

Propositional density (PD) is the exception to this rule, and it is defined as the ratio between the number of propositions (P) and concepts (C).

2.3 Data analysis

2.3.1 Descriptive statistics and t-test

Box-plot graphs were used to show the statistical parameters (average, median, lower quartile, upper quartile, sample minimum, and sample maximum) obtained to describe each variable considered in our analysis. These graphs allow checking the central tendency, the symmetry, and the dispersion of a data set at a glance (Cohen & Lea, 2004).

A comparison of the averages for each variable was made by using t-test. It usually determines whether two means are significantly different from each other, and allowed to compare the values that describe Cmaps obtained during classes 1 and 15. The larger the t value of the test, the more likely the test is statistically significant and the compared averages are different (Cramer & Howitt, 2004).

2.3.2 Cluster analysis: multivariate statistics

A data matrix X(69,3) containing only selected parameters (PD, MIC, and FMC) was used to carry out hierarchical cluster analysis (HCA). HCA is an exploratory statistical procedure to classify class 1 and class 15 Cmaps into groups according to their similarities (Kaufman & Rousseeuw, 2005). Statistica 8 (StatSoft, Tulsa, OK, USA) was the software chosen to run this analysis. The graphical output of HCA (dendrogram) will be used to discuss the results in the next section.

3 Results and discussion

3.1 Descriptive statistics: comparing the results for class-1 and class-15 Cmaps

The 69 Cmaps were analyzed considering the 8-parameter set indicated in Table 2. Average and standard deviation for the subsets of Cmaps obtained in class 1 (n=14, collaborative) and class 15 (n=55, individual) were presented for comparative purposes (Table 3).

Parameter	Class-1 Collaborative Cmaps (n=14)	Class-15 Individual Cmaps (n=55)	Calculated t	Are the averages significantly different?[a]
PD	1.0±0.2	1.7±0.4	6.32	Yes
OIC	1.8±1.4	1.4±1.0	1.23	No
OFC	2.8±1.2	2.3±1.3	1.30	No
IFC	5.3±2.0	6.4±1.7	2.09	No
IC	7.1±1.1	7.7±1.4	1.49	No
MIC	2.1±1.3	4.9±1.6	6.05	Yes
FC	8.1±1.4	8.6±1.1	1.43	No
MFC	1.6±1.3	5.3±1.7	7.58	Yes

[a]Critical value: $t(p=0.0005, df=\infty) = 3.29$.

Table 3. Averages for the parameters considered during the structural analysis of Cmaps and the results obtained by using t-test

The average comparison using t-test allowed the identification of 5 parameters that did not change significantly. The calculated t for OIC (t=1.23), OFC (t=1.30), IFC (t=2.09), IC (t=1.49), and FC (t=1.43) was lower than the critical value (3.29). Therefore, the increase of the students' experience with concept mapping did not affect them. On the other hand, the parameters PD (t=6.32), MIC (t=6.05), and MFC (t=7.58) showed calculated t values far above than the critical value (3.29). Moreover, the averages for PD, MIC, and MFC were higher for class 15 than class 1 Cmaps. It was possible to identify these parameters as potential "fingerprints" to monitor the students' proficiency on concept mapping, since in our case, they increased when the students' experience with concept mapping increases. For this reason, the following discussions consider only propositional density (PD), initial concepts with multiple propositions (MIC), and multiple final concepts with multiple propositions (MFC).

3.2 Box-plots for the devised categories

Figure 3 presents the box-plots obtained for the selected parameters (PD, MIC, and FMC) after comparing the averages for class 1 and class 15 Cmaps. The average (Av) and the median (Me) of propositional density (PD) were Av=1.0±0.2/Me=1.0 and Av=1.7±0.4/Me=1.7, for class-1 and class-15 Cmaps, respectively. This noteworthy increase indicated that trained students can include more propositions during the set up of concept maps than beginners. This fact may be related to the new way of expression that is imposed by concept mapping. Students need to express his/her ideas, by using concise, clear, and precise elements (linking phrase, initial and final concepts) to construct meaning. Readers and writers may "unlearn" the manner they usually elaborate their oral/written discourses to work with a more interconnected net of keywords (concepts) and brief explanations (linking phrases).

The transition from linear thinking and understanding (texts) towards a systemic perspective (Cmaps) was confirmed by the increase of average and median for MIC and MFC (class 1: Av=0.2±0.1/Me=0.2; class 15: Av=0.5 ±0.2/ Me=0.6 for both). Moreover, MIC and MFC may indicate the occurrence of progressive differentiation (MIC) and integrative reconciliation (MFC) during the Cmap elaboration. As these processes are related to meaningful learning (Ausubel, 2000), we can assume that MIC and MFC may be useful parameters to verify students' options into the root-meaningful learning continuum.

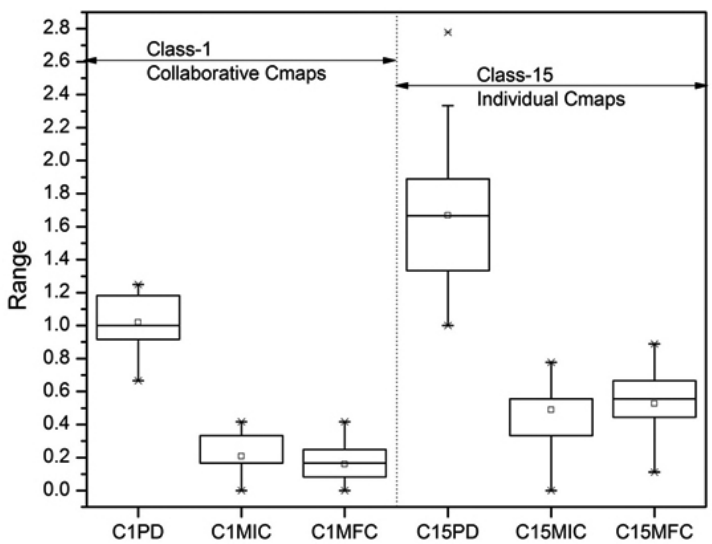

Figure 3. Box-plot obtained for selected parameters of the structural analysis (PD, MIC, and MFC) to highlight the differences between class-1 and class-15 Cmaps

3.3 Finding patterns through cluster analysis

Hierarchical cluster analysis (HCA) was used as a multivariate exploratory analysis (Kaufman & Rousseeuw, 2005). A subset of the data obtained through structural analysis was organized into as a matrix X(69,3). Only PD, MIC, and MFC were considered for clustering purposes due to the results shown in Table 3.

Several methods for measuring distances were tested in order to achieve the most meaningful clusters, considering our expertise on concept mapping. Distances between objects (Cmaps) were measured using single, complete, and Manhattan (or City-Block) methods. Distances between clusters were measured using Euclidean and Ward's methods. The dendrogram presented in Figure 4 was obtained using Manhattan and Ward's methods for measuring distances between object and clusters, respectively. An arbitrary value corresponding to 15% of de maximum distance between 2 objects (Cmaps) was chosen to define the clusters to be compared. The dotted line in Figure 4 helps to identify 5 clusters (I-V).

Despite the descriptive parameters presented in Table 3 for all Cmaps (n=69), the average for the selected variables (PD, MIC, and MFC) was calculated for characterizing each cluster. The differences between clusters can be discussed from the data presented in Table 4. The use of black dots was an option to make easier to compare the distinguishable features of each cluster. In spite of all quantitative manipulations of the data, the qualitative approach prevails when HCA is used. Therefore, the information in Table 4 is presented in a friendly graphical manner.

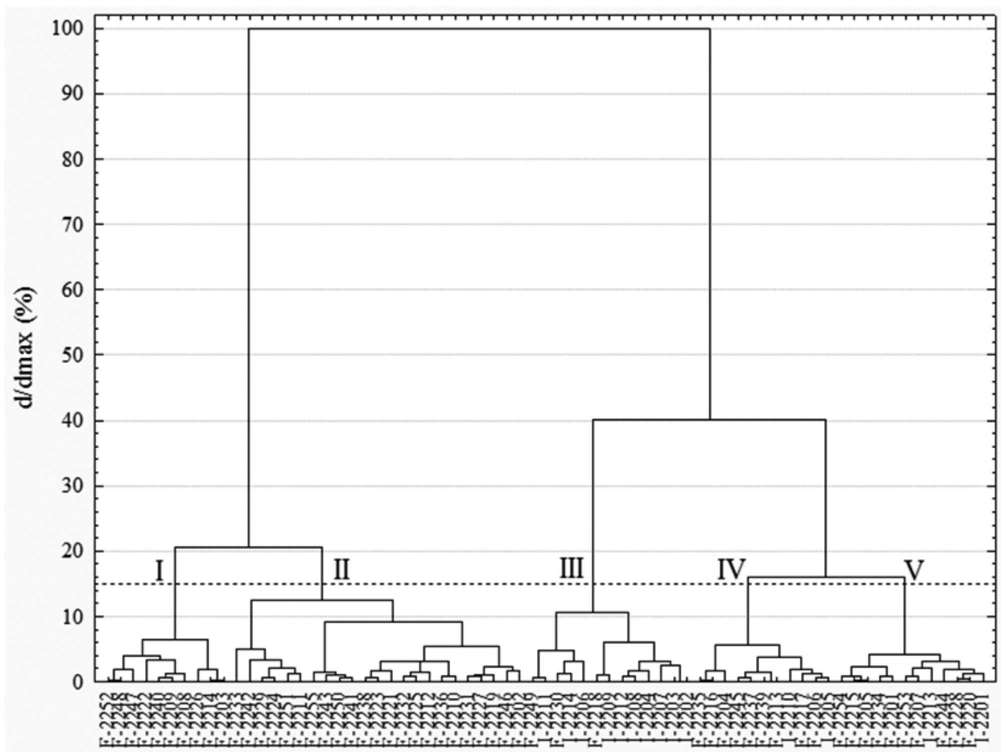

Figure 4. Dendrogram obtained by HCA considering the matrix X(69,3). Selected parameters for running HCA: Manhattan (or City Block) distance to measure between CMaps distance; Ward distance to measure between clusters distance. An arbitrary value of 15% of the maximum distance (dmax) was selected to characterize the clusters (I-V). Notation in the dendrogram: F-type Cmaps are from class 15 (n=55); I-type Cmaps are from class 1 (n=14)

	Cluster ID	**I**	**II**	**III**	**IV**	**V**
	# of CMs	*10*	*23*	*13*	*11*	*12*
Parameters for structural analysis	*PD*[a]	•••	••••	•	••	••
	MIC[a]	•••	••••	•	••	••
	MFC[a]	•••	••••	•	••	••
	IFC	••	••••	••	•••	•
	IC	••	••••	••	•••	•
	FC	••	••••	••	•••	•
	OIC	•••	•	•••	••	••••
	OFC	•••	•	•••	••	••••

[a]Parameters considering for clustering purposes.

Table 4. Differences between clusters explained by the parameters used for structural analysis. The number of black dots indicates the relative average magnitude for each cluster. Highest values (••••), lowest values (•), intermediate values (••) or (•••)

The highest average values (••••) for PD, MIC, and MFC described cluster II, which contains the best Cmaps in structural terms. It should be highlighted that cluster II presents only Cmaps made by experienced students (class 15). The opposite could be verified for cluster III, which contains the lowest average values (•) for PD, MIC, and MFC. This cluster concentrates the majority of the Cmaps made by naïve mappers.

Intermediate values (••/•••) were useful to characterize clusters I (••• for PD, MIC, and MFC), IV, and V (•• for PD, MIC, and MFC). Cluster I is similar to cluster II and it also contains only Cmaps from class 15. The structural qualities found in Cmaps classified in clusters I and II are only achieved by well-trained mappers. On the other hand, the properties found in Cmaps assigned in cluster III characterize beginners with no or little experience in concept mapping. Clusters IV and V were mainly formed by Cmaps from class 15 that are not so good in comparison to their counterparts classified in clusters I and II. Four Cmaps from class 1 were allocated in clusters IV and V, suggesting they should be better than the pattern found for the majority of the other Cmaps prepared by the beginners.

A summary of the clusters' description from the selected parameters (PD, MIC, and MFC) makes clear the dependency between mappers' proficiency and the Cmap structure.

- Highest values (••••) for PD, MIC, and MFC: cluster II, which contains the best class 15 Cmaps.
- Lowest values (•) for PD, MIC, and MFC: cluster III, which contains the majority of class 1 Cmaps.
- Intermediate values (•••) for PD, MIC, and MFC: cluster I, which is similar to cluster II and contains only class 15 Cmaps.
- Intermediate values (••) for PD, MIC, and MFC: clusters IV and V, which contains ordinary class 15 Cmaps and 4 class 1 Cmaps.

4 Summary

Mappers' proficiency on concept mapping is critical to achieve all promised rewards we hope to find when using Cmaps to represent and share knowledge. Training sessions for beginner mappers should be explored to overcome the naïve use of this technique. This work showed the relationship between mappers' proficiency and the structure of the Cmaps. The quantitative structural analysis was proposed to describe the fine Cmap morphology. We found that propositional density (PD), initial concepts with multiple propositions (MIC), and final concepts with multiple propositions (MFC) can be use to estimate the proficiency level of the mappers. The increase of these parameters indicates a better understanding about how to set up a concept map. Despite this work being explored Cmaps about science teaching, we believe that the quantitative structural analysis has a wide range of applications, including both educational and corporative uses.

5 Acknowledgements

The authors thank CAPES (3555/09-7), CNPq (553710/2006-0) and FAPESP (06/03083-0, 08/04709-6) for funding our research group. S. N. Silva Jr and J. G. Romano Jr thanks the Pró-Reitoria de Graduação da Universidade de São Paulo and the Secretaria Estadual da Educação for their scholarships. We also thank Profa Dra Káthia Maria Honorio for all insightful discussion about exploratory multivariate analysis and hierarchical cluster analysis.

References

Ausubel, D. P. (2000). The acquisition and retention of knowledge: a cognitive view. Dordrecht: Kluwer Academic Publishers.

BouJaoude, S., & Attieh, M. (2008). The effect of using concept maps as study tools on achievement in chemistry. Eurasia Journal of Mathematics, Science & Technology Education, 4(3), 233-246.

Cañas, A. J., & Novak, J. D. (2006). Re-examining the foundations for effective use of concept maps. In A. J. Cañas & J. D. Novak (Eds.), Proceedings of the Second International Conference on Concept Mapping (pp. 247-255). San Jose, Costa Rica: Universidad de Costa Rica.

Cañas, A. J., & Novak, J. D. (2006). Re-examining the foundations for effective use of concept maps. In A. J. Cañas

& J. D. Novak (Eds.), Proceedings of the Second International Conference on Concept Mapping (pp. 247-255). San Jose, Costa Rica: Universidad de Costa Rica.

Coffey, J. W., Eskridge, T. C., & Sanches, D. P. (2004). A case study in knowledge elicitation for institutional memory preservation using concept maps. In A. J. Cañas & J. D. Novak (Eds.), Proceedings of the First International Conference on Concept Mapping (pp. 274-281). Pamplona, Spain: Universidad Pública de Navarra.

Coffey, J. W. (2006). In the heat of the moment: strategies, tactics, and lessons learned regarding interactive knowledge modeling with concept maps. In A. J. Cañas & J. D. Novak (Eds.), Proceedings of the Second International Conference on Concept Mapping (pp. 137-145). San Jose, Costa Rica: Universidad de Costa Rica.

Cohen, B. H., & Lea, R. B. (2004). Essentials of statistics for the social and behavioral sciences. Hoboken: John Wiley & Sons.

Correia P. R. M., Infante-Malachias M. E., & Godoy C. E. C. (2008). From theory to practice: the foundations for training students to make collaborative concept maps. In: A. J. Cañas, P. Reiska, M. Ahlberg, & J. D. Novak (Eds.), Proceedings of the Third International Conference on Concept Mapping (pp. 146-153). Tallin (Estonia) & Helsinki (Finland): OÜ Vali Press.

Correia P. R. M., Valle, B. X., Dazzani, M., & Infante-Malachias, M. E. (2010). The importance of scientific literacy in fostering education for sustainability: theoretical considerations and preliminary findings from a Brazilian experience. Journal of Cleaner Production, 18(7), 678–685.

Cramer, D. & Howitt, D. (2004). The SAGE dictionary of statistics: a practical resource for students in the social sciences. London: SAGE Publications Ltd.

Fourie, L. C. H., van der Westhuizen, T. (2008). The value and use of concept maps in the alignment of strategic intent. In: A. J. Cañas, P. Reiska, M. Ahlberg, & J. D. Novak (Eds.), Proceedings of the Third International Conference on Concept Mapping (pp. 399-406). Tallin (Estonia) & Helsinki (Finland): OÜ Vali Press.

Gerstner, S., & Bogner, F. X. (2009). Concept map structure, gender and teaching methods: an investigation of students' science learning. Educational Research, 51(4), 425-438.

Kaufman, L., & Rousseeuw, P. J. (2005). Find groups in data: an introduction to cluster analysis. Hoboken: John Wiley & Sons.

Kinchin, I. M., Hay, D. B., & Adams, A. (2000). How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development. Educational Research, 42(1), 43-57.

Kinchin, I. M., & Alias, M. (2005). Exploiting variations in concept map morphology as a lesson-planning tool for trainee teachers in higher education. Journal of In-service Education, 31(3), 569-591.

Kyrö, P., & Niskanen, V. A. (2008). A concept-map modeling approach to business planning in a computer environment. In: A. J. Cañas, P. Reiska, M. Ahlberg, & J. D. Novak (Eds.), Proceedings of the Third International Conference on Concept Mapping (pp. 262-269). Tallin (Estonia) & Helsinki (Finland): OÜ Vali Press.

Novak, J. D. (2010). Learning, creating, and using knowledge: concept maps as facilitative tools in schools and corporations. 2nd Ed. New York: Routledge.

Santos, W. L. P. (2007). Scientific literacy: a freirean perspective as a radical view of humanistic science education. Science Education, 93(2), 361-382.

Torres, P. L., & Marriott, R. C. V. (2009). Handbook of research on collaborative learning using concept mapping. Hershey: IGI Global.