Representing Domain Knowledge by Concept Maps: How to Validate Them?

Dietrich Albert, Christina Steiner

Representing Domain Knowledge by Concept Maps: How to Validate Them?

Dietrich Albert and Christina M. Steiner
Cognitive Science Section – Department of Psychology – University of Graz
Universitätsplatz 2 / III, 8010 Graz, Austria
{dietrich.albert, chr.steiner}@uni-graz.at

Abstract: For one and the same domain several alternative concept maps may exist, originating from different world views or purposes. Some of these concept maps may be valid, however not all of them. Thus, strategies for empirically and objectively validating concept maps in the respective context are necessary. We outline two methodological approaches for empirically validating concept maps, one for giving evidence of content validity and one for application validity. One procedure is to validate a given concept map with concept maps systematically generated by others. As a second method we suggest to observe the behaviour and the performance in a relevant situational context as a validation criterion. In this scope, a method for predicting persons' problem solving behaviour by using a given concept map is outlined. In general, the purpose and ultimate use of a given concept map has to be taken into consideration for choosing a validation procedure and interpreting its results.

Keywords: concept map, semantic net, ontology, domain knowledge, validation, content validity, application validity

1 Introduction

Concept maps (semantic nets) provide a valuable tool for organising and presenting knowledge within an ontological framework. They serve a variety of purposes, especially in educational contexts (for an overview see e.g. Coffey, Carnot, Feltovich, Feltovich, Hoffman, Cañas, & Novak, 2003; Novak, 2001), but e.g. also in the context of hypertexts (as tools for hypermedia design and navigation). Most commonly, a concept map is depicted by a labelled, directed graph with the vertices (nodes) representing concepts of a domain and the directed and labelled edges (arcs) representing the relationships that exist between those concepts. Figure 1 shows an example of a concept map, representing what a concept map is.

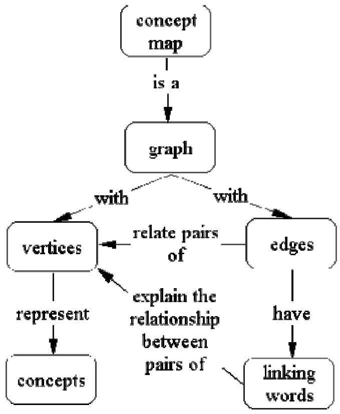

Figure 1: Concept map describing what a concept map is (adapted from Ruiz-Primo, 2000)

The combination of two concepts and the link relating them forms a meaningful statement and therefore constitutes a proposition. Hence, another way of representing a concept map is a list of propositions. Furthermore, the information contained in a concept map may also be presented in form of a matrix, with the set of concepts labelling the columns and rows and the relations specified in the cells of the matrix.

Mathematically, a concept map can be defined as follows. A concept map is a directed graph consisting of a finite, non-empty set $C = \{c_1, ..., c_n\}$ of nodes, representing concepts (or concept labels) and a finite, non-empty set A of arcs, representing the relations between concepts. Every arc in A is an ordered pair of concepts from the set C. The relations in a concept map are labelled; each relation label i defines a binary relation R_i on the set of concepts. A concept c_p is in relation R_i to another concept c_q, i.e. $R_i(c_p,c_q)$, if and only if there exists an arc in the concept map with the label i and with an arrowhead pointing to the second node of the ordered pair (c_p,c_q). Each arc is an element of one relation R_i, the arc set A is the union of all relations R_i.

Representations of semantic knowledge as described by superordination (e.g. Heller 1994, 2000), mind maps (Buzan & Buzan, 1996), cognitive maps (e.g. Ackermann, Eden, & Cropper, 1992) constitute special cases of a concept map, each featuring only one relation (is a, related to, leads to). Please note, that semantic knowledge representations known as e.g. semantic nets (e.g. Fisher, Wandersee, & Moody, 2000) and knowledge maps (e.g. O'Donnell, Dansereau, & Hall, 2002) are also in line with the mathematical definition of concept maps given above. Thus, although in the sequel we concentrate on concept maps, the given considerations and explanations actually also hold for those methods of knowledge representation.

For generating concept maps, aside from the traditional methods using paper and pencil or other devices, a variety of software exists, e.g. Inspiration (www.inspiration.com), Hozo (http://www.hozo.jp), CmapTools (http://cmap.ihmc.us), or Protégé-2000 (http://protege.stanford.edu). Utilising such software provides significant support in creating, manipulating, storing, and reusing concept maps.

A concept map may represent a knowledge domain, e.g. for presenting learning material in an e-learning environment. A concept map is also able to represent personal knowledge, e.g. when a student is asked to generate a concept map specifying his/her personal understanding of a specific domain. Here, we focus on concept maps as domain knowledge representations.

Basically, two approaches can be distinguished in building a concept map for a particular domain, a normative and a descriptive one. When holding the view of a normative approach, it is assumed that there exists only one concept map of complete consensus for the domain. Contrary, according to the descriptive approach, it is assumed that there may exist alternative concept maps for a specific domain. In other words, for a given domain there may be different, but not arbitrary concept maps. Generally, the descriptive approach seems to be more reasonable instead of demanding 'the one correct' concept map. This is, because a concept map necessarily entails some sort of world view or opinion regarding a particular domain. Hence, concept maps representing the same domain may differ from each other. Moreover, differences may occur by reason of the intended purpose and ultimate use of concept maps.

One general aim in building concept maps is to obtain at least one of possibly several valid concept maps of a domain. A crucial question in this context therefore refers to the validity of a given concept map. As a result, a framework for evaluating the adequacy of a concept map or of different proposals for a concept map of a specific domain should be provided. This field of research, however, still lacks efficient strategies. In the present paper we elaborate on the question of validating concept maps representing domain knowledge. Before outlining two types of validity that we distinguish and two approaches for examining them, the topic of reliability of concept maps is shortly discussed.

2 Reliability of Concept Maps

Before addressing the validity of a given concept map, the issue of its reliability should be regarded. Test-retest reliability can easily be determined by having a concept map for the same domain generated again. Let us assume an expert of a specific knowledge domain who constructed a concept map for this domain. By asking the respective person in a different point of time to generate a concept map for the same domain again, the reliability of the concept map can be examined. For this, it is assumed, that there is no indication of a change in knowledge or understanding of the person. If both concept maps correspond to each other, they reliably represent (the domain or at least) the understanding of the respective person regarding the domain. If the expert generates a different concept map each time, though, it is obvious that no indication exists that one of them will be a reliable model of the knowledge in question.

A given concept map can be described by attributing specific characteristics or features to the respective map. This is the case when scoring a concept map according to a particular scoring system (e.g. Novak & Gowin, 1984; Ruiz-Primo, 2000) that considers different map characteristics. Some of these features can be objectively and clearly determined, e.g. number of concepts, number of propositions. Contrary, there are also map characteristics that have to be determined by subjective ratings, e.g. proposition accuracy. The reliability of such ratings can be examined by using again the test-retest method. Another procedure refers to the interrater-reliability, i.e. the consistency of ratings assigned to a concept map (e.g. Ruiz-Primo, 2000). This aspect of reliability can be examined when having two or more judges that independently from each other score a given concept map.

Assuming again an expert that generated a concept map representing a particular domain, parallel-forms reliability can be examined by asking to construct a concept map in alternative forms of representation (e.g. as a directed graph and as a list of propositions). A reliable concept map should represent the same model of knowledge (for the domain or at least) for the respective person, regardless in which representation format it is generated.

3 Validity of Concept Maps

Having considered the reliability of a concept map, the aspect of validity can be addressed. A concept map representing domain knowledge constitutes a model of a part of the current knowledge about the world for a given domain in a given context. Such a model may serve e.g. for presenting learning material, for predicting problem solving performance etc. Thus, it is important to ensure, that a concept map is well founded and valid.

Subjective evidence for the correctness of a concept map is not enough for making a statement regarding validity. Even the principle of consensus with respect to the correctness of a concept map does not suffice. In fact, objective and empirical criteria are needed for giving evidence of the validity of a given concept map.

The validity of a concept map can be understood from two perspectives. On the one hand, it may be examined whether a concept map serves the purpose for that it has been designed. This aspect of validity could be denoted as 'application validity' of a concept map. It refers to the practical usability and usefulness of a concept map. Therefore, different kinds of intended application will require different means of validating a concept map.

Before examining the quality of a concept map by applying it for the purpose it has been generated, though, evidence needs to be given regarding its validity of the knowledge in question. This means, it has to be determined whether the concept map constitutes a valid model of a part of the current knowledge about the world. This aspect of validity we call 'content validity'. Content validity is an important issue, as the evaluation of the content of concept maps is critical for using them. Of course, the evaluation whether a concept map adequately reflects the respective knowledge will also need to take into account its intended purpose and ultimate use.

In the following sections we concentrate on both, content and application validity.

3.1 Content Validity

As a valuable approach for giving evidence of the content validity of a given concept map, we suggest to take empirically collected concept maps representing personal knowledge (in the sequel denoted as 'criterion maps') as a criterion. Concept mapping tasks are an appropriate way for eliciting persons' understanding of a domain. Thus, for gaining information regarding the content validity of a given concept map representing a particular domain, it may be compared with empirically gathered concept maps of individuals of different knowledge level, including experts. For this, the similarity between the given concept map and the criterion maps is examined.

This can be done by investigating the conformance of propositions. To this end, for instance 2x2 tables can be established for comparing the given concept map with a criterion map, detailing how many propositions are contained either in both or in only one of them. Based on this, in correspondence to Ruiz-Primo (2000) for example a so-called 'convergence score' (*Con*) can be calculated. Utilising this score for validating a given concept map representing domain knowledge, it would express the proportion of propositions in the given concept map out of the total number of propositions in one criterion map. In this case, the respective score is derived by $Con = x / y_k$, whereas x denotes the number of propositions in the given concept map that are also contained in the criterion map k and y_k denotes the total number of propositions in the criterion map k. Other statistical methods for determining the similarity between given concept map and criterion concept maps can be found e.g. in Goodman and Kruskal (1979) or Tversky (1977).

Of course, such similarity measures can not only be determined for propositions, but also only for relations or concepts, respectively, if only those special parts of a given concept map are to be validated.

If in sum a high similarity between the given concept map and the criterion maps can be determined, this indicates, by definition, the content validity of the given concept map.

A given concept map may not necessarily have to be validated as a whole. Possibly only a part of a concept map has to be validated, e.g. the contained concepts or relations, a substructure of the map etc. In this case either only parts of the criterion maps are used for validation or only partial criterion maps are collected. Thus, depending on the validation objectives, a set of different procedures for posing a concept mapping task is available. Applying the 'construct-a-map' or 'map creation' method (e.g. Ruiz-Primo, 2000), individuals are asked to generate a concept map concerning a specific knowledge domain from scratch – either by providing concepts and/or relations or not. This may be done by drawing the concept map by hand, by arranging note-cards, or of course by using suitable software as mentioned in the introductory section. The 'fill-in-the-map' or 'map completion' technique, (e.g. Schau, Mattern, Zeilik, Teague, & Weber, 2001) is characterised by providing a concept map of a particular domain to individuals, where all or some of the concepts and/or relations have been left out. The blanks then have to be filled in. The software CMap Pro[1] (http://www.uni-saarland.de/~su11pshb/forsch/cmap.html), for example, features the possibility of creating fill-in maps that can afterwards be presented and filled in directly on the computer display. One further alternative of a concept mapping task makes use of relatedness ratings between pairs of concepts (e.g. Schau et al., 2001). For this, an individual is asked to rate the degree of relatedness between pairs of previously defined concepts on a numerical scale. Through applying a mathematical algorithm, e.g. by using the Pathfinder software (http://interlinkinc.net/Pathfinder.html), the relatedness ratings of an individual can then be visually represented in form of a graph. Based on the resulting graph an individual may additionally be asked to label the arcs between concept nodes (Shavelson, Ruiz-Primo, & Wiley, in press) in order to get a concept map. Another kind of concept mapping task consists in presenting the propositions of a concept map as a correct-incorrect discrimination task, which should additionally include distractor items and confidence ratings (Steiner, 2004). A version of this method - however without the possibility of including distractors or confidence ratings - has also been implemented in the software CMap Pro, mentioned before, namely for assessing pre-knowledge.

The different techniques of posing a concept mapping task vary considerably in the extend of constraints imposed and information provided to individuals. All of them are conceivable for empirically collecting criterion maps, when the content validity of a given concept map, or parts of it, is to be examined. Of course, each method has special characteristics with respect to validation. Which method is most suitable will depend very much on the particular validation issue and its requirements, e.g. whether a whole given concept map or only specific parts of it are to be validated. Furthermore e.g. the cognitive and the time demands will influence the method for obtaining criterion maps. In any case the purpose and context has to be taken into account when collecting criterion maps for validation.

3.2 Application Validity

For examining the application validity of a given concept map, we propose utilising situational performance as a criterion of validation. With situational performance we mean behaviour in real-world situations which does not consist in performing a concept mapping task. This could for example be problem solving, answering questions, or even social behaviour in given situations. It seems natural and obvious, that an individual's personal understanding of a domain is reflected in his/her behaviour and performance in given situations, at least as declarative knowledge is involved. Situational performance is therefore useable as a

[1] The development of this software unfortunately has stopped.

criterion for examining the validity of a concept map. For validating a given concept map, a kind of situational performance has to be chosen, that is related to the purpose and intended application of the given concept map. Due to different purposes (e.g. presenting learning material, describing social skills) that are intended for a given concept map, different kinds of situational performance (e.g. problem solving, behaviour in social situations) will be appropriate for validation. Thus, specific situational performance scores or profiles of individuals of different knowledge level, including experts, are collected and constitute the validation criterion.

Performance scores such as number of correct answers in e.g. multiple choice tests (e.g. Rice, Ryan, & Samson 1998; Ruiz-Primo, 2000; Schau et al., 2001), standardised tests (e.g. Rice et al., 1998), or problem solving tasks (e.g. Steiner, 2004) have already been applied as criteria for validating personal concept maps. In sum, it could be shown that measures of situational performance are suitable for validating concept maps representing personal knowledge or at least, that individual concept maps and performance are somehow positively related, indicating application validity in this special case. Of course these performance measures can be used for validating any given concept map. However, measures like test scores constitute only summary validation criteria.

Instead of or in addition to these criteria, it would be desirable to apply more sophisticated performance profiles as criteria for validation. This is why we suggest to take Knowledge Space Theory (Albert & Lukas, 1999; Doignon & Falmagne, 1999; Falmagne, Koppen, Villano, Doignon, & Johannesen, 1990) as a suitable framework for validation. We propose a validation approach that utilises Knowledge Space Theory for predicting problem solving behaviour, i.e. for predicting answer patterns, based on a given concept map. The given concept map then can be validated by comparing the predicted answer patterns with empirically obtained answer patterns. In the following section a short introduction into the basic notions of Knowledge Space Theory is given. Subsequently, the validation approach mentioned above will be outlined in more detail.

3.2.1 Basic Notions of Knowledge Space Theory

Knowledge Space Theory provides a formal model for structuring a domain of knowledge and for representing the knowledge of individuals based on prerequisite relationships. A domain of knowledge is characterised by a finite, non-empty set Q of problems. The knowledge state K of a learner is represented by the subset of problems that he/she is capable of solving.

Due to mutual dependencies among the problems of a domain, from the correct solution of certain problems the mastery of other problems can be surmised. Such relationships between problems are captured by the so-called surmise relation. The surmise relation S is a binary relation on the set Q of problems, that is reflexive and transitive. Two problems a and b are in a surmise relation, i.e. $(a,b) \in S$, whenever from a correct solution to problem b the mastery of problem a can be surmised. In other words, problem a is a prerequisite problem for problem b.

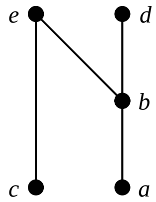

Figure 2**:** Example of a Hasse diagram illustrating a surmise relation on a knowledge domain $Q = \{a, b, c, d, e\}$ (adapted from Falmagne et al., 1990)

A surmise relation can be depicted by a so-called Hasse diagram (see Figure 2 for an example). In such a diagram descending sequences of line segments indicate a surmise relationship. According to the surmise relation illustrated in Figure 2, from a correct solution to problem b the correct solution to problem a can be surmised, while the mastery of problem e implies correct answers to problems a, b, and c. The surmise relation forms a quasi-order on the set Q of problems and thus restricts the number of possible knowledge states (i.e. subsets of problems) that are expected to be observable. The collection of all possible knowledge states, including the empty state $\emptyset$ and the whole set Q, constitutes the so-called knowledge structure $\mathcal{K}$. The knowledge structure $\mathcal{K}$ corresponding to the surmise relation shown in Figure 2 is given by

$$\mathcal{K} = \{\emptyset, \{a\}, \{c\}, \{a, c\}, \{a, b\}, \{a, b, c\}, \{a, b, d\}, \{a, b, c, e\}, \{a, b, c, d\}, Q\}.$$

3.2.2 Validation Approach

Let us assume a concept map representing the declarative knowledge of a particular domain, for which application validity is to be examined. As an appropriate measure of situational performance to be applied as validation criterion collecting problem solving patterns has been chosen. To this end, a set of typical and representative problems of the domain is selected. Each problem is identified with a subset of propositions of the given concept map, representing those elements of declarative knowledge that are required for solving the respective problem. In other words, each problem is mapped on the given concept map, by assigning the subset of propositions required for mastering the respective problem. Each proposition can be considered as an atomic skill or competency in the sense of the approaches of Doignon, Düntsch and Gediga, and Korossy (for references see Albert & Lukas, 1999; Doignon & Falmagne, 1999). Most likely, the subsets of propositions assigned to the problems will overlap. Based on the representation of the problems by subsets of the given concept map, dependencies between problems in terms of a surmise relation can be derived. This could be done e.g. by set inclusion, i.e. if the representation of a problem a in the concept map is a subset of that of problem b, then problem a is a prerequisite for problem *b*. From the dependencies between the problems derived in this way, possible knowledge states can be identified and a knowledge structure can be established. This means that specific answer patterns, that are expected to be observable, out of all possible subsets of items ($2^{|Q|}$) can be predicted. In the following, the described procedure is illustrated by an example. Assuming a given concept map representing the domain in question, containing a set P of propositions (whereas a proposition $(c_p,c_q) \in R$ is abbreviated by p_j), that is given by

$$P = \{p_1, p_2,\ldots, p_j,\ldots, p_7, p_8\}.$$

Let furthermore the selected set Q of problems representing the same domain be

$$Q = \{a, b, c, d\}.$$

Assume, the mapping m of the problems on the concept map is specified as follows:

$$m(a) = \{p_1, p_3\}$$
$$m(b) = \{p_2, p_3, p_4\}$$
$$m(c) = \{p_1, p_2, p_3, p_4, p_5\}$$
$$m(d) = \{\{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8\}$$

From the assignment of propositions to the problems given above it can be seen e.g., that problem a can be identified with the propositions p_1 and p_3, i.e. p_1 and p_3 represent the declarative knowledge required for solving problem a. By applying the principle of set-inclusion, dependencies between the problems can be derived:

$$m(a) \subseteq m(c) \Rightarrow (a,c) \in S$$
$$m(a) \subseteq m(d) \Rightarrow (a,d) \in S$$
$$m(b) \subseteq m(c) \Rightarrow (b,c) \in S$$
$$m(b) \subseteq m(d) \Rightarrow (b,d) \in S$$
$$m(c) \subseteq m(d) \Rightarrow (c,d) \in S$$

From the derived dependencies a surmise relation can be deduced, which is illustrated in Figure 3.

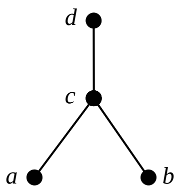

Figure 3: Surmise relation based on the mapping of the set Q of problems on the given concept map

The knowledge structure that is induced by the surmise relation shown in Figure 3 is given by

$$\mathcal{K} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{a, b, c\}, \{a, b, c, d\}\}.$$

In this way, based on the mapping of the problems on the concept map, six possible knowledge states have been identified. These knowledge states constitute answer patterns that are expected to be observed, provided that the given concept map adequately represents the domain.

The next step in the validation approach is therefore to collect empirical answer patterns, by posing the problems to individuals of different knowledge level, including experts. In this way, it can be investigated empirically whether the observed answer patterns correspond to the identified and predicted knowledge states. This would be done e.g. by using a discrepancy index describing the similarity between the knowledge structure and the set of answer patterns (e.g. Doignon & Falmagne, 1999, Chapter 12). As the knowledge structure has been established based on mapping the problems on the given concept map, the empirically obtained answer patterns serve as a criterion for validity. If the empirical answer patterns correspond well to the predicted knowledge states, the given concept map can be considered to be valid – provided that both, the chosen set of problems as well as the sample of persons are adequate and representative.

4 Conclusions

In general, it is not sufficient to subjectively evaluate a concept map, judging it as valid by subjective evidence or uncontrolled consensus. Actually, there is an urgent need of objective, empirical measures and criteria for giving evidence of the validity of a concept map (semantic net). Efforts for empirically and objectively validating concept maps have proven reasonable and promising.

We distinguish two types of validity, content validity and application validity. The former refers to the adequate representation of the domain modelled, the latter refers to the practical usefulness of a concept map. We suggested and outlined two approaches for empirically validating a given concept map, one for giving evidence of content validity and one for application validity. Both approaches constitute useful and valuable procedures. For building up a coherent picture about the validity of a given concept map, both approaches can complement each other, providing a comprehensive view from different perspectives. Of course, the outlined approaches are not the only possibilities for validating concept maps. There are also other ways that constitute suitable validation approaches, as e.g. consulting the published literature of the given knowledge domain for examining content validity.

One special feature of the suggested validation approaches is that not only experts in a given field, for which a concept map is to be validated, are queried or consulted. In fact, the proposed procedures rather involve individuals of different knowledge level, who will possibly afterwards perform occupational tasks based on validated concept maps, e.g. developing a curriculum.

In general, when validating a given concept map, the purpose and ultimate use (e.g. predicting problem solving behaviour) needs to be taken into account. This issue is very important, as the content and structure of a concept map may highly depend on its intended purpose. Furthermore, validation efforts of concept maps also need to consider which aspects are intended to be validated (e.g. the whole concept map, the concepts, the relations, substructures of the given concept map). Another critical point is to choose appropriate statistical measures, depending on the particular criterion for validity and its constraints.

It should be mentioned that to date there is also little attention paid to the reliability of concept maps. Actually, before focusing on the validity of a given concept map, some considerations should be dedicated to reliability. Does the given or generated model of knowledge reliably represent the knowledge in question? The question of reliability or at least the circumstances under which the concept map originated, should be taken into account when the validity of a concept map is addressed.

For a given knowledge domain there might be several alternative concept maps that validly represent the respective domain. Most likely, such alternative concept maps will match in parts, i.e. with respect to concepts and relations, or even in whole substructures. An interesting research topic in this context is that of merging concept maps. Merging means a process that tries to integrate the information from two or more concept maps (or other forms of semantic knowledge representation) into a single one. How this can be done best is a subject of graph theory and ongoing research (e.g. Noy & Musen, 2000; Stefanutti, Albert, & Hockemeyer, 2000).

Summarising, the issue of validating concept maps and other forms of semantic knowledge representation, such as semantic nets, knowledge maps etc., is an emerging field of research. Having available well-founded and valid concept maps is critical for their effective use, be it in the classroom, in the context of hypertexts etc. Publishing a concept map, reusing an existing map for building a new one, or implementing an application that relies on concept maps written by others or even yourself without first evaluating it would be very unwise (Gómez-Pérez, 2001). The methodological considerations presented in this paper provide interesting starting points for practical validation efforts.

Acknowledgements

Part of the work reported in this paper was financially supported by the European Commission through the iClass project (Contract Number 507922) within the FP6-IST programme.

We wish to thank Jürgen Heller for his helpful comments on an earlier version of this paper.

References

Ackermann, F., Eden, C., & Cropper, S,. (1992) Cognitive Mapping: Getting Started with Cognitive Mapping. *Proceedings of the 7th Young Operational Research Conference, University of Warwick, April 1992* (pp. 65-82).

Albert, D. & Lukas, J. (1999). *Knowledge Spaces: Theories, Empirical Research, and Applications.* Mahwah, NJ: Lawrence Erlbaum Associates.

Buzan, T. & Buzan, B. (1996). *The Mind Map Book: How to Use Radiant Thinking to Maximize Your Brain's Untapped Potential*. New York: Plume.

Coffey, J.W., Carnot, M.J., Feltovich, P.J., Feltovich, J., Hoffman, R.R., Cañas, A.J., & Novak, J.D. (2003). A *Summary of Literature Pertaining to the Use of Concept Mapping Techniques and Technologies for Education and Performance Support.* Technical Report submitted to the Chief of Naval Education and Training. Pensacola, FL. Retrieved June 6, 2005, from Institute for Human and Machine Cognition (IHMC): http://www.ihmc.us/users/acanas/Publications/ConceptMapLitReview/IHMC%20Literature%20Review%20on%20Concept%20Mapping.pdf

Doignon, J.-P. & Falmagne, J.-C. (1999). *Knowledge Spaces*. Berlin: Springer.

Falmagne, J.-C., Koppen, M., Villano, M., Doignon, J.-P., & Johannesen, L. (1990). Introduction to Knowledge Spaces: How to Build, Test, and Search Them. *Psychological Review, 97*, 201-224.

Fisher, K.M., Wandersee, J.H., & Moody, D. (2000). *Mapping biology knowledge.* Dordrecht: Kluwer.

Gómez-Pérez, A. (2001). Evaluation of Ontologies. *International Journal of Intelligent Systems, 16*, 391-409.

Goodman, L.A. & Kruskal, W.H. (1979). *Measures of association for cross classifications.* New York: Springer.

Heller, J. (1994). Semantic Structures. In D. Albert (Ed.), *Knowledge Structures* (pp. 117–149). Mahwah, NJ: Lawrence Erlbaum Associates.

Heller, J. (2000). Representation and Assessment of Individual Semantic Knowledge. *Methods of Psychological Research - Online, 5.* Retrieved August 4, 2005, from http://www.mpr-online.de/issue10/art1/heller.pdf

Novak, J.D. (2001). *The Theory Underlying Concept Maps and How To Construct Them.* Retrieved May 10, 2005, from Institute for Human and Machine Cognition (IHMC): http://cmap.coginst.uwf.edu/info/printer.html

Novak, J.D. & Gowin, B.D. (1984). *Learning how to learn*. New York: Campridge University Press.

Noy, N.F. & Musen, M.A. (2000). PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI-2000),* Austin, TX.

O'Donnell, A.M., Dansereau, D.F., & Hall, R.H. (2002). Knowledge Maps as Scaffolds for Cognitive Processing. *Educational Psychology Review, 14*, 71-86.

Rice, D.C., Ryan, J.M, & Samson, S.M. (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. (2000). On the Use Of Concept Maps As An Assessment Tool in Science: What We Have Learned so Far. *Revista Electrónica de Investigación Educativa, 2*. Retrieved May 30, 2005, from http://redie.uabc.mx/vol2no1/ contents-ruizpri.html

Schau, C., Mattern, N., Zeilik, M., Teague, K. W. & Weber, R. (2001). Select-and-Fill-In Concept Map Scores as a Measure of Students' Connected Understanding of Science. *Educational and Psychological Measurement, 61*, 136-158.

Shavelson, R.J., Ruiz-Primo, M.A., & Wiley, E. (In Press). Windows into the Mind. *International Journal of Higher Education.* Retrieved May 30, 2005, from Stanford Education Assessment Laboratory: http://www.stanford.edu/dept/SUSE/SEAL/Reports_Papers/Windows%20into%20the%20Mind_8_4_03_Final.doc

Stefanutti, L., Hockemeyer, C., & Albert, D. (2003). Derivation of Knowledge Structures for Distributed Learning Objects. In T. Dimitrakos, P. Ritrovato & S. Salerno (Eds.), *3rd International LeGE-WG Workshop: GRID Infrastructure to Support Future Technology Enhanced Learning*. Wiltshire: British Computer Society.

Steiner, C.M. (2004). *Beziehung zwischen Concept Maps und Kompetenzstrukturen.* Unpublished diploma thesis: University of Graz.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.