Implicit Concept Mapping: A Computerized Tool for Knowledge Assessment in Undergraduate Psychology

Eugene Aidman and Janine Ward¹

Abstract

The study continued validating the computerised method of implicit concept mapping (Aidman & Egan, 1998), while extending it from assessing the map’s structural properties to content-based expert evaluation. The on-line concept mapping task (Aidman & Egan, 1998) was modified to elicit similarity / contrast judgements for a set of basic personality concepts, in a group of 65 introductory psychology students. The resulting individual concept proximity matrices were scored for complexity and internal consistency, as well as individually factor- and cluster analysed. Hierarchical cluster tree and un-rotated factorial representations were generated for each individual map. Students were asked to interpret their own cluster trees and factor plots by naming the clusters and factor axes (a brief statement accompanying the name was allowed). Three independent experts (lecturers in the subject) rated the clarity and accuracy of these interpretations, as well as the soundness of cluster trees and the factorial representations themselves. These data were compared with the overall grade the students had

¹ Authors’ notes: Eugene Aidman and Janine Ward (School of Behavioural & Social Sciences & Humanities, University of Ballarat, Victoria, Australia).

Earlier versions of this paper were presented at the Annual Conference of the Australian Psychological Society, Sydney (NSW), September 1996, the International Conference on the Application of Psychology to Quality of Learning and Teaching, Hong Kong, June, 1998, and “Tertiary Teaching and Learning: Dealing with Diversity” Conference, Darwin (NT), 9-12 July 2001. The authors acknowledge the support of Australian Research Council (Small ARC Grant to the first author). We thank Dennis Else and Ray Over for their vision and critique of this project.

All correspondence concerning this article should be sent to Dr Eugene Aidman at eugene.aidman@dsto.defence.gov.au
obtained in the Introductory Psychology course immediately prior to participating in the study. Internal consistency and complexity of students' implicit concept maps produced only crude associations with their course achievement. The expert ratings of cluster- and factorial representations of the maps, especially expert ratings of students' own interpretations of their maps, showed a more refined association with course achievement, thus confirming that structural assessment of the implicit maps needs to be complemented by their content-based evaluation in order to achieve a more accurate estimate of the learner's level of expertise in the subject.

1. Introduction

Concept mapping as a visual method of knowledge representation, is widely regarded not only as an effective learning technique (Novak, 1990) but also as a promising tool for knowledge evaluation and assessment (c/f Aidman & Egan, 1998; Novak, 1993; Wilson, 1994). Since pioneering efforts of Novak and his colleagues (c/f., Novak, 1977, 1981, 1990; Novak & Gowin, 1984) concept mapping has established its utility in both exploratory (Trochim, 1989; Linton, 1989; Trochim, Cook & Setze, 1994) and evaluative (Neimeyer, 1989; Beyerlein, Beyerlein & Markley, 1991; Markham, Mintzes & Jones, 1994; Wilson, 1994) representation of knowledge. This method is now widely used in science teaching (Sandoval, 1995), as well as in a range of contexts such as teacher education (Winitzky, Kauchack & Kelly, 1994), program evaluation and planning (Trochim, 1989; Trochim, Cook & Setze, 1994), instructional design (Barenholtz & Tamir, 1992; Starr & Krajcik, 1990; Edmondson, 1994, 1995) and evaluation of conceptual change (Beyerbach & Smith, 1990; Wallace & Mintzes, 1990; Songer & Mintzes, 1994; Trowbridge & Wandersee, 1994).

Concept mapping has proven particularly useful in helping students learn and teachers teach. In particular, concept maps have been found useful for structuring and integrating students' knowledge (Fisher, 1990; Reader & Hammond, 1993), for "describing student's evolving constructions of knowledge" (Beyerbach & Smith, 1990, p. 961), measuring changes in knowledge structure and the identification and modification of students' limited or inappropriate conceptual frameworks (Novak, 1990). Furthermore, the use of concept mapping as a learning tool has been found to reduce the students' anxiety towards learning a topic in the classroom (Jegede, Alaiyemola & Okebukola, 1990). By visually revealing the ways learners organize the knowledge that they hold,
concept mapping helps the learner to identify misconceptions and a poor understanding; and encourages new interpretations of old ideas, as well as other forms of creative thinking (Roth, 1994). It has been suggested that concept maps may be sensitive to changes in the students' knowledge structure (Novak & Wandersee, 1990) and may have the potential to reflect the individual's evolving understanding of a given area of learning (Trowbridge & Wandersee, 1993).

Studying individual differences in conceptual structures, thus, offers promise as a means for evaluating and assessing knowledge. Concept maps have been found to be capable of differentiating between knowledge structures of advanced biology majors from beginning non-majors in the domain of zoology (Markham, Mintzes & Jones, 1994). Concept mapping has also shown the ability to capture the differences in cognitive structure between experts and novices in the area of art history (Beyerlein, Beyerlein and Markley, 1991). The latter study showed that both measures of knowledge breadth and depth derived from conceptual mapping procedures were strongly correlated with membership in the expert and novice group. These findings confirmed that experts not only possess a greater amount of knowledge (compared to non-experts), but have it remarkably better organised.

It is not surprising that concept mapping has emerged as a significant development in computer-aided learning (Trapp, Reader & Hammond, 1992). The relationship between concept mapping and academic assessment has, however, received relatively little attention (Beyerlein, Beyerlein & Markley, 1991; Wilson, 1994). Concept mapping measures have so far shown only low to moderate correlations with achievement measures such as course grades and scholastic aptitude tests (see Novak, Gowin & Johansen, 1983). However, concept maps have been demonstrated to discriminate between contrast groups of learners, such as novices and advanced learners (Wilson, 1994; Markham, Mintzes & Jones, 1994). Concept mapping has been found sensitive to differences in knowledge that traditional multiple choice tests were unable to detect (Wallace & Mintzes, 1990). Goldsmith and Johnson (1990) point out that “the basic problem with conventional assessments is that they fail to recognise that much is not most disciplinary knowledge is based on an understanding of relationships among concepts”. Structural representations such as concept maps may capture this property better than conventional assessment methods (Markham, Mintzes & Jones, 1994).

It is widely recognised that "the manner in which students are tested determines more than any other factor ... the ways in which students learn" (Fisher, 1990, p. 1015).
The problem with conventional instructional techniques and assessment methods is that they make readily foreseeable demands on learners (Rothkopf, 1988). It is not surprising that Novak (1990) found learning patterns of Cornell University students not reaching beyond “essentially rote learning most of the time” (p. 942), which provides a vivid evidence that learning strategies, learning goals and assessment methods are intimately linked. The fact that CM has been successfully used as a learning strategy implies there are particularly suited learning goals. When the learning goal is to build internal connections among conceptual elements and apply what is learned to new situations, then strategies, such as CM, which involve integration, organisation and inference, should lead to more successful outcomes. Novak’s (1990) suggested that CM may lead to more meaningful (and less rote) learning because it emphasises concepts, principles and propositions. A judicious utilisation of CM in knowledge assessment would therefore be able to enhance these outcomes by stimulating the use of CM as a leaning strategy.

Novak, Gowin and Johnasen (1983) reported that although concept mapping scores did not correlate with common assessment techniques they did with tests requiring higher levels of understanding. This implies that concept mapping either tap abilities that are not well measured by common assessment techniques, or abilities that are not assessed at all. Markham, Mintzes and Jones (1994) submit that these findings illustrate the inadequacies of traditional assessment techniques in differentiating between a deeper level, meaningful understanding of a given domain, and a superficial, rote learned one.

Given the propensity of concept maps to be a more sensitive conceptual measure, and one which appears to discriminate more effectively between rote and meaningful learning, its potential as an adjunct or alternative mode of assessment has been well recognised (Markham, Mintzes & Jones, 1994). CM has been able to reveal considerable changes in the complexity and propositional structure of learners’ knowledge base even after brief periods of instruction (Wallace & Mintzes, 1990). It was also successfully used to monitor evolving constructions of knowledge over longer periods of time (Beyerbach & Smith, 1990).

Overall, CM appears to provide a viable assessment framework, tapping into the deeper layers of learners’ knowledge. Not only have concept maps proven effective in monitoring conceptual change over periods of time, they have also shown a potential to ascertain an individual’s comprehension of the subject, and consequently to differentiate (at least approximately) between individuals of different mastery / achievement levels.
in physics (Larkin, 1983; Pankratius, 1990), chemistry (Gussarsky, Gorodetsky, 1988, 1990; Wilson, 1994), human movement (Blais, 1993), art history (Beyerlein, Beyerlein and Markley, 1991) zoology (Markham, Mintzes & Jones, 1994) and psychology (Durso & Coggins, 1990). However, these studies examined only a broad-band sensitivity of CM (e.g., expert-novice differentiation). The present study specifically addresses the issue of how accurate a CM assessment can be in differentiating the grades and gradations of academic achievement.

2. **Quantitative evaluation of concept maps.**

Two distinct approaches have been used to quantify concept maps. The predominant system stems from the original work of Novak (1981) and entails scoring the number of concepts the person links, levels of hierarchy, valid relationships, branching, cross-links and examples (Novak & Gowin, 1984). Several weighting schemes that have been used to quantify these elements (Stuart, 1985; Markham, Mintzes & Jones, 1994), all involving manual scoring by domain experts and in some cases control for inter-rater reliability. Measures established by this method yield reasonable differentiation between expert and novice learners (Gobbo & Chi, 1986; Beyerlein, Beyerlein & Markley, 1991) and change as understanding of a subject increases (Trowbridge & Wandersee, 1993; Wallace & Mintzes, 1990). The measures also provided some reflection of learner achievement in physics (Pankratius, 1990), chemistry (Wilson, 1994), human movement (Blais, 1993), and biology (Songer & Mintzes, 1994). The accuracy of this methodology, however, has not gone much beyond crude classifications such as between “high” and “low” achievers in secondary school chemistry (Wilson, 1994) or between non-major freshmen and advanced graduates in biology (Markham, Mintzes & Jones, 1994). Often individual maps are collapsed into group matrices (Wilson, 1994), which leads to further loss of discriminating power.

An alternative approach involves direct learner rating of inter-concept similarities and subsequent multivariate decomposition of the concept proximity matrix resulting in a reconstruction of underlying cognitive maps (Aidman & Egan, 1998; Trochim, 1989). The map is then interpreted, usually by the learners themselves (c/f. Trochim, Cook & Setze, 1994). The concept proximity matrix can be decomposed by multidimensional scaling (Trochim, 1989; Keith, 1989), cluster or factor analysis (Aidman & Egan, 1998). This method is essentially based on Kelly’s (1955) long established *personal construct* framework and its repertory grid methodology (Fransella & Bannister,
The methodology was recently extended to include other methods of multivariate decomposition such as pathfinder analysis (Schvaneldt, 1991) which led to successful hybrid approaches (Wilson, 1994) combining both Trochim’s and Novak’s ideas of map quantification.

Aidman and Egan (1998) extended Trochim’s methodology to evaluating individual conceptualisations. Using an on-line knowledge mapping software (DCS-4; Burmistrov & Shmeliov, 1991, 1992), Aidman and Egan (1998) conducted on-line reconstruction of individual implicit cognitive maps from learner-derived concept proximity data and examined their configural properties, such as their complexity, internal consistency and similarity to an expert “model” map, in relation to learner achievement in the subject. The utility of these structural measures in differentiating students at several levels of achievement will be analysed. The study that follows is designed to further validate this method of reconstructing individual concept maps.

The current study employed mapping a segment of conceptual knowledge about personality in the context of an introductory undergraduate course in psychology. It was hypothesised that both structural / configural and content-evaluation measures, derived from the reconstructed concept maps, will vary as a function of the students’ level of achievement in the subject.

3. Method

Subjects. 65 first year students (modal age 19 years) volunteered to participate in the study. All student participants received course credit in Learning, Perception and Cognition - an introductory course in cognitive psychology they were all enrolled in at an Australian regional university.

Materials. The pre-selected set of eleven concepts, derived from the sub-domain of personality, represented a well-defined segment of the prescribed study content (Weiten, 1992). All participants performed a mapping task with this eleven-item concept list (see Figure 1). Mapping a pre-determined set of concepts is widely recognised as a viable technique (c/f., Trapp, Reader & Hammond, 1992; Wilson, 1994), particularly because of the standardisation it affords.

Procedure. The mapping task, which was administered via a knowledge mapping software package DCS-4 (Burmistrov & Shmeliov, 1992), required participants to make a series of similarity and contrast judgments between concepts. Each concept took its
turn at the top of the list (header concept) and the task was to select from the remaining list not only the concept most similar to the header, but also two concepts most in contrast with it (an example can be seen in Figure 1). The process was repeated with subsequent header concepts selected from the list until the list is exhausted. The resulting proximity matrix was then analysed to generate spatial-dimensional and hierarchical cluster tree representations of the individual’s map. The same proximity matrix was further analysed to compute global structural characteristics of the map, such as cognitive complexity and internal consistency.

<table>
<thead>
<tr>
<th>Header Concept</th>
<th>The Unconscious</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most similar to Header</td>
<td>Concept List</td>
</tr>
<tr>
<td>Archetype</td>
<td>Locus of Control</td>
</tr>
<tr>
<td></td>
<td>Reinforcement</td>
</tr>
<tr>
<td></td>
<td>Projection</td>
</tr>
<tr>
<td></td>
<td>Conditioning</td>
</tr>
<tr>
<td></td>
<td>Self-actualisation</td>
</tr>
<tr>
<td></td>
<td>Super-Ego</td>
</tr>
<tr>
<td></td>
<td>Inferiority complex</td>
</tr>
</tbody>
</table>

**Figure 1:** A computer screen dump showing the process of concept comparison: Moving concepts left if similar to Header and right if contrasting (explanation in text).
Data analysis. The on-line data collection and preliminary analysis were handled by the DCS-4 program (Burmistrov & Shmeliov, 1991). The proximity data were derived for each individual subject in the form of an \( n \times n \) non-symmetrical square matrix \( A_n = \{a_{ij}\} \), where \( n \) = number of concepts mapped. The cells values were assigned as follows: \( a_{ij} = 1 \) if \( j \)th concept was judged as similar to \( i \)th concept, \( a_{ij} = -1 \) if \( j \)th concept was judged as contrasting \( i \)th concept, and \( a_{ij} = 0 \) if neither similarity nor contrast among \( i \)th and \( j \)th concepts was recorded. An example of \( A_n \) matrix can be seen in Appendix A. Note that \( A_n \) is not a distance matrix in itself: it is non-symmetrical and as such, can not be factorised. In order to produce a correlation-like, factorisable metric, each \( A_n \) was modified into a normalised matrix of all pairwise scalar products of its rows in a procedure developed and tested by Burmistrov and Shmeliov (1992). The resulting symmetrical matrices \( S_n \) were factor analysed using Harman’s principal components search with a standard \( \textit{eigenvalue greater then 1.0} \) criterion for factor selection. This procedure usually returns between two and five components. Un-rotated factor loadings for the first two components were transformed to a two-dimensional representation, with all other pair-wise projections of significant components available on request and adding to the visualisation of the original matrix’s factorial structure, thus representing spatial relationships between the mapped concepts (see Figure 2).

Due to individual differences in the format of cognitive aggregation, spatial-factorial representation may not be the most appropriate form of representation to everyone: a substantial proportion of the population prefers taxonomy-based representations (Aidman, 2001). In order to cater for this variation, proximity matrices were further analysed with Johnson’s hierarchical clustering algorithm, which resulted in a cluster-tree representation for each individual map, segmented into comparable levels of similarity among the mapped concepts (see Figure 3).

Cognitive complexity of a representation is typically evaluated by an estimate of its independent elements, or by the degree to which a construct system is broken down (Bieri, 1986). In our case, the cognitive complexity of the map was estimated with a \textit{differentiation} index, which reflects the number of independent dimensions to which the map can be reduced (Burmistrov & Shmeliov, 1992). The index is computed as follows:
\[ D = 1 - \sum_{i=1}^{n} \sum_{j=1}^{i-1} \frac{abs(S_{ij})}{(m \times n(n-1)/2)} \]

Where \( n \) = number of concepts mapped, \( m \) = number of non-zero cells in each row of matrix \( \{a_{ij}\} \), and \( S_{ij} \) is the scalar product of the \( i \)th and \( j \)th rows of the matrix:

\[ S_{ij} = \sum_{k=1}^{n} a_{ik} \times a_{jk} \]

\( D \) is effectively a measure of dimensionality of a proximity matrix, with its values ranging from 0 to +1.0. \( D \) has been reported to increase as the subject distinguished more independent properties in the given knowledge domain (Burmistrov & Shmeliov, 1992), with low \( D \) values indicating a simpler cognitive structure.

Internal consistency of the map was estimated by a measure of symmetry of the proximity matrix:

\[ C = 1 - \sum_{i=1}^{n} \sum_{j=1}^{i-1} \frac{SIG_{ij}}{(n(n-1)/2)} \]

where \( SIG_{ij} = 1 \) if \( \text{sign}(P_{ij}) = \text{sign}(Q_{ij}) \), and \( SIG_{ij} = 0 \) if \( \text{sign}(P_{ij}) \neq \text{sign}(Q_{ij}) \), where \( P_{ij} \) is the scalar product of the \( i \)th and \( j \)th rows and \( Q_{ij} \) is the scalar product of \( i \)th and \( j \)th columns of matrix \( \{a_{ij}\} \). \( C \) values also range within \([0,1]\) interval, reaching +1.0 for a symmetrical matrix representing a perfectly consistent set of judgements, and diminishing with the increased asymmetry of the matrix.

---

**Figure 2:** A spatial-factorial representation of a student implicit map of personality concepts.
Hierarchical cluster tree and un-rotated factorial representations were generated for each individual map (see Figures 2 and 3). Both factorial and cluster representations were evaluated by three independent experts (lecturers in the subject), using a rating scale from 1 (inadequate dimensions / cluster structure) to 5 (fully appropriate structure). Students were also asked to interpret their own cluster trees and factor plots by naming the clusters and factor axes (a brief statement accompanying the name was allowed). The experts rated these interpretations as well, on a scale from 1 (inadequate content labelling) to 5 (fully competent). These data were compared with the overall grade the students had obtained in the Introductory Psychology course immediately prior to participating in the study.

Figure 3: Hierarchical cluster tree representation of an implicit map of personality concepts.
4. **Results and Discussion**

The results showed that, while the complexity of student's implicit maps of personality terms was not associated with course grade (a one-way ANOVA showed no significant differences between grade groups on Differentiation index, $F(5, 59) = 0.147, p > 0.05$), the internal consistency of these implicit maps was associated with the grades ($F(5, 59) = 3.16, p < 0.05$). However, post-hoc analysis revealed that map consistency could only differentiate top students (A grade) from the rest of the field (see Figure 4).

The expert ratings of both cluster tree and factorial representations of the students' implicit maps -- and, particularly, expert ratings of their interpretations by map owner - showed a more refined association with the overall grade ($F(5, 59) = 4.55, p < 0.01$, $F(5, 59) = 5.21, p < 0.01$ and $F(5, 59) = 7.60, p < 0.001$ respectively, see Figure 5 for illustration). Leaving aside the controversial issue of how representative the grades are of the students' knowledge of the subject, these data suggest that structural assessment of implicit maps needs to be complemented by content-based evaluation in order to achieve a more accurate estimate of the learner's level of expertise in the subject. Implicit mapping may be used as a reliable source of material for such content-based evaluation - the cluster- and factorial representations of these maps.

Different forms of content-based evaluation were not equally effective, according to our data. In fact, the tree expert ratings – ratings of students' factor structure, cluster structure and own map interpretation - showed quite distinct patterns of utility. In particular, factor structure ratings showed a greater overall range of assessment and better discrimination of extreme grade groups (A and F) from the middle grades (B, C and D). Cluster structure ratings, on the other hand, produced an overall smaller range of assessment that only separated the combined upper grades group from the group of marginally passing and failing students. Further, own map interpretation ratings not only separated the top grades from the middle, but also were quite sensitive to differences at the lower end of the range, producing significant differences between those with credit, pass and fail grades. These findings are not indicative of a clear supremacy of any of the tree forms of content-based evaluation – rather, each of the three expert ratings may play a differential role in the overall assessment of learners’ competence.

Aidman and Egan's (1998) results showed that implicit mapping is sensitive to individual differences in both domain specific and more generic characteristics of learners' knowledge. This was evidenced in concept mapping indices correlating with both the
students' overall marks on a formal academic test covering the mapped content among other topics and the sub-test scores reflecting exclusively the mapped content. Interestingly enough, the latter association is weaker than the former, indicating that concept mapping reflects generic learner characteristics better than their domain specific knowledge.

![Diagram showing differentiation and consistency scores for different grades](image)

**Figure 4:** Structural assessment of implicit maps of personality concepts: Map complexity (differentiation) and internal consistency scores for students with different levels of achievement (grades) in the course.

In a broader sense, the utility of concept mapping measures in the assessment of individual's knowledge beyond the content from which those measures are derived is not only entirely consistent with conventional academic assessment practices, but is necessary for any assessment method if it is to find widespread use. In this sense the results demonstrated an acceptable level of concurrent validity of concept mapping against conventional academic assessment. At the same time, however, as was expected, the association found between concept mapping and traditional assessment is incomplete (evidenced by low to moderate correlations of various concept mapping indices and standard multiple-choice test scores). This incomplete association indicates that apart from the shared variance in learners' knowledge reflected in both concept mapping and multiple-choice test scores, there is unique variance in learners' performance that can be explained only by concept mapping. Concept mapping thus appears to tap either abilities that are not well measured by common assessment techniques or abilities that are not assessed at all. Although insufficient to reach a firm conclusion, the results indicate
that concept mapping reflects generic learner characteristics better than their domain specific knowledge, which is consistent with a widely held view that concept mapping is a more sensitive form of knowledge assessment (Markham, Mintzes & Jones, 1994), and one that tends to discriminate more effectively between rote and meaningfully learned information.

![Figure 5: Expert ratings of cluster structure, factor structure, and own map interpretation (personality concepts) for students with different levels of achievement (grades) in the course.](image)

**Figure 5:** Expert ratings of cluster structure, factor structure, and own map interpretation (personality concepts) for students with different levels of achievement (grades) in the course.

Although earlier reports (Novak, Gowin & Johansen, 1983) indicated that concept mapping scores did not correlate with common assessment techniques, such as multiple choice questions, associations were evident with tests requiring higher levels of understanding. Markham, Mintzes and Jones (1994) submit that these findings illustrate the inadequacies of traditional assessment techniques in differentiating between a deeper level, meaningful understanding of a given domain, and a superficial, rote-learned one. Our results support this contention, though partially and indirectly, in showing stronger concept mapping correlations with more generic learner appraisal (which is more likely to be influenced by overall quality of learning practices than a narrow domain achievement). The findings thus provide further support to Novak's (1990) claim that concept mapping should lead to more meaningful (and less rote) learning. This support, however, comes from a distinctly different perspective: when the assessment outcomes are tied more to a broader knowledge base than to memorizing material in a narrow domain, the learner is quietly but inevitably encouraged to adopt more sound learning strategies, with or without concept mapping among them. Thus, concept mapping assessment becomes instrumental in overall enhancement of learner practices whether or not they include concept mapping as a learning strategy.
5. References


