Abstract

'Repertory grid analysis' was used as a way of constructing representations of learners conceptions of living things; and finding a common structure or understanding; this was described in previous work. 'Concept mapping' has been stated as an appropriate assessment procedure in science curricula in a number of countries. Previous relevant published work describing the benefits of 'concept mapping' as a learning, teaching, and assessing tool is large. However, the existence of a large literature is not justification in itself, and in this work we provide a critique of the current literature. Problems with 'concept mapping' are simply ignored. However, it is thought that 'repertory grid analysis' does overcome some of the recognized problems of 'concept mapping' such as: i) comparison of concept maps between students or between students and educators: in repertory grid this becomes not only possible but a reliable method - this has proven to be one of the more problematic areas in traditional 'concept mapping'; ii) quantification of concept mapping - integer scoring of whole concept maps in 'repertory grid analysis' is eliminated since the arranged structure of the graph produced holds significance for the conceptual structure. More research needs to be done in repertory grid analysis and its implications and applications in curricular research have yet to be fully explored.

Key words: concept mapping, conceptual frameworks, repertory grid analysis.

Introduction

Knowledge of how concepts are classified and categorized, terms which are differentiated by Estes (1996) as they become assimilated, leads to the assumption that if the learning environment, i.e., 'schooling', is adapted to maximize the acquisition and ordering of concepts, such an adaptation should consider the ordering processes. In a similar thread, scientists order concepts of living things in diagrams which are akin to concept maps, therefore one way to learn, and assess the learning of, the 'correct' system of ordering living things would be through such diagrams. Such a diagram would allow the educator to find out easily if the categories formed are formal-biological or folk-biological or somewhere in-between, in the case of living things. McCloughlin and Matthews (2009, 2010) did outline three studies into school-based classifications:

(i) How one group classified five terms of equines: horse, donkey, mule, zebra and pony; and the term 'goat' as an 'out-group'.
(ii) How computer generated drawings of equines were classified generally,
using living vs. extinct, unfamiliar vs. familiar exemplars over the post-primary school age-range

(iii) How formal drawings of dicotyledonous plants were classified with two professional biologists as an out-group.

The technique employed by McCloughlin and Matthews (2009, 2010) was an adaptation of repertory grid analysis; however, it is thought that diagrams which represent the ordering of concepts may relate to each other, and this is the subject of this present paper. The aim of this work is to examine the repertory grid technique itself and concept mapping in more detail in relation to constructing a learner’s conceptual framework of concrete representations of abstractions (i.e., concretistic concepts), with concepts of living things being used as an example. It is part of an on-going project into the applications of Repertory Grid Analysis (RGA) in science education research.

What Is Concept Mapping?

Concept mapping is a way to represent concepts in a diagrammatic manner showing the links between concepts and their components. The term ‘mapping’ is applied because it is thought that the spatial arrangement of concepts denotes a significance, a measureable ‘mental distance’ (Wandersee, 1990), except that in concept mapping the distance are not ‘measured’. Concept maps are typically representations of simple propositions of the ‘subject-link-predicate’ type as shown in Figure 1. Concept mapping was defined by Pankratius and Keith (1987) as “a two-dimensional hierarchical representation of concepts which indicates the relationship between selected concepts”. Fisher (2000) places concept mapping as one of many types of knowledge mapping. She states that it was invented, and extensively researched, by Joseph D. Novak (Wandersee, 1990) and his research team. Concept mapping has been employed in a large number of areas within the area of education. Some of these are: as a learning tool (Horton et al., 1993; Roth & Roychoudhury, 1993); measuring achievement (Pankratius, 1990); reduction of learners’ anxiety (Okebukola & Jegede, 1989); integration within curricular design (Nicoll, Francisico, & Nakleth, 2001); assessment (Atkinson & Bannister, 1998; Buldu & Buldu, 2010; Laffey & Singer, 1997; Liu, 1994; McClure & Bell, 1990; Oliver & Raubenheimer, 2006; Robinson, 1999; Ruiz-Primo & Shavelson, 1996); and self assessment (Stow, 1997); autobiographical reflections (Lim, 2011); collaborative learning (G. Wilson, Nash, & Earl, 2010); conceptual growth (Miller et al., 2009) – categories which are not mutually exclusive.

![Socrates](image)

Figure 1: A simple single proposition as the most basic form of a concept map.
What is Repertory Grid Analysis?

George Kelly (Kelly, 1992) formulated the seminal work on repertory grid analysis (RGA) which was initially grounded in clinical psychology – and indeed, it is this domain which has made the most extensive use of this technique. RGA was devised to identify, and investigate the relationships between, a person’s mental constructions. Subjects rate, score or rank the relevance, importance or similarity of a characteristic to a list of entities, i.e. concepts – this list comprises the so-called repertory or repertoire. This produces a grid or matrix of $n \times m$ form of $n$ columns and $m$ rows (Figure 2.). In classical repertory grid analysis, the subject chooses those characteristics which are pertinent to the test item derived by, or supplied to, the subject. The process of ranking elements against concepts, called ‘constructs’, is a constructive process, in that a representation of the subjects mental ordering has been made explicit in a stepwise fashion.

![Figure 2: DISPLAY: A simple matrix of ratings of four elements (people) against three constructs (descriptions).](image)

Physically, a mathematical entity has been produced but importantly, it has been constructed by the subject indirectly and it represents that part of the ‘mindscape’, what Kelly called ‘psychological space’, concerned with these concepts. The basis of RGA is the elicitation of a grid or matrix of integers where each number is a score of relevance of a feature to a range of constructs. Once a numerical matrix is formed, it can be subjected to many kinds of statistical analysis (Fransella, Bell, & Bannister, 2004) such as cluster analysis (Figure 3.), principal components analysis (Slater, 1964, 1977) (Figure 4.). In such a graph, ‘Phaedrus’ is a ‘corruptor of youth’ while ‘Socrates’ is characterised as ‘philosopher’ and to a lesser degree a ‘questioner of ideas’. ‘Lysias’ is an ambiguous person being near the origin, and ‘Euthyphro’ is a ‘Greek’ ‘man’. However, all four people had all three traits, but to varying degrees of perception. Note also that the constructs, in blue, are bipolar, but that the characteristics at either end of the pole are not simple ‘opposites’ but rather alternatives from the repertoire of roles this set of people may be perceived to have.
The result of this analysis is the production of a set of co-ordinates derived from principal component loadings, typically, the first and second loadings are utilised to plot the position of the constructs on a Cartesian plane. The analyses produce graphs that are representations of a person’s set of constructs with the co-ordinates of the constructs plotted as points in two or more dimensions. The selection of only the first and second principal component loadings is again the subject of further work (McCloughlin and Matthews, in preparation).

The mathematical techniques used in RGA can be used to analyse the structures of learners’ concepts in science, and other curriculum subject areas, even though such concepts do not show all the characteristics required of the constructs traditionally employed in RGA. However, there are few studies that apply RGA to the study of concepts commonly found in the sciences (Bezzi, 1996; Fetherstonhaugh, 1994; Winer & Vazquez-Abad, 1995). However, its main use in educational research to date has been in studying the perceptions held by educators of their work, the children in their care or themselves ‘in role’; and/of their students of them, their environment and the material they must learn (Artiles & Trent, 1990; Cuniliffe, 1994; Derry & Potts, 1998; Fang, 1996; Kubrusly, 1984; Menmuir & Christie, 1999; Munby, 1984; Nespor, 1985; Owens, 1988; Shaw, 1992; Solas, 1992; Tisher, 1983; Walker & Kleine, 1985; Watson, 1994; Williams, Pack, & Khisty, 1997)
What Then Is Conceptual Framework Mapping?

“Conceptual framework” is a generic term to include techniques for representing concepts and the relationships between them pictorially. In this paper, we consider the use of repertory grid analysis and concept mapping as techniques of conceptual framework mapping and review its relationship with ‘classical’ concept mapping and briefly with modern biological techniques.

Problem of Research

Kinchin (2001) asked the still highly pertinent question that if concept mapping is so helpful to learning science then “why aren’t we all doing it?” Kinchin (2000) believed that there has been a reluctance in the science education researcher community to address problems with concept mapping in the face of purported wide ranging advantages and benefits ascribed to concept mapping. The simple truth is that in spite of the positive literature, teachers of primary, secondary and tertiary science do not use concept mapping as much as they could. The purpose of this paper is to deal with the sort of problems that have been outlined by a number of researchers.

Kinchin (2001) describes the main problem a teacher faces when having administered a concept mapping task to a class group: what do you do with them now? How does one compare 20-30 (or more) concept maps with each other and with the teachers’ ‘authoritative’ conception? Taking one step back further, an even simpler question is how to compare only two concepts that one person constructs, and thus examine the progress of the learner. A key problem of attempting to employ concept maps in education has been to devise protocols to apply a score or ‘mark’ in order to introduce an element of summative assessment. Indeed, early attempts in applying a score to concept maps involved a summative approach (Novak & Gowin, 1984). The constructivist agenda is contrary to impressionistic scoring of test items and issue of validity and reliability are relevant also.

Can Concept Maps Be Assessed?

In this section, a number of methods of scoring concept maps are reviewed before RGA is examined in more detail. The following methods all attempt to decompose concept maps into their components and score the components based on some kind of weighting. The scores are then analysed according to an algorithm, usually a summation. Novakian concept mapping (Markham, Mintzes, & Jones, 1994) has undergone a number of adaptations by different workers (Martin, Mintzes, & Clavijo, 2000), some of whom developed a system for scoring structural complexity and examining propositional validity.

\[
S_{cm} = n_c + n_r + n_h + n_b + \sqrt{\frac{n_r}{n_c}}
\]  

(i)

Where:
- \(n_c\) is the number of points for concepts, one for each non-redundant concept,
- \(n_r\) is the number of points for relationships, one point for each valid, scientifically-acceptable proposition,
- \(n_h\) is the number of points for hierarchy, five points for each level of hierarchy,
- \(n_b\) is the number of points for branching, one point for the first branching and three points for each additional branching.
In a simpler approach the score is the sum of individual scores allotted to:

(i) the number of links \textit{per se}: score of 1 per link
(ii) the number of levels or nestings in a hierarchy: score of 5 per level
(iii) the number of cross links: score of 2 or 10 per cross link depending on which version of the protocol one is using, and
(iv) the numbers of examples: score of 1 per example.

This is summarised in the following equality:

\[ S_{cm} = n_l + n_n + n_c + n_e \]  

Where:
- \( S_{cm} \) is the concept map score
- \( n_l \) is the number of links, each carrying a weighting of 1
- \( n_n \) is the number of nestings, each carrying a weighting of 5
- \( n_c \) is the number of cross links, each carrying a weighting of 2 or 10
- \( n_e \) is the number of examples, each carrying a weighting of 1.

Whereas these summative approaches have the benefit of simplicity, there are two main causes of concern in its application. Firstly, although the concept map may provide an impression of structure and a complex structure may appear to indicate a rich understanding of the content relating to the concept map, the score merely reflects a measure of the complexity rather than the quality or richness. Unfortunately it does not provide any comment on the structure of the ideas expressed. Secondly, as mentioned before, students who have been trained to construct concept maps would acquire the scoring rules and aim to maximise their scores through cross-links. The rebuff to the second point might be to say that only the relevant cross-links would be scored. However, the issue of relevance is fraught with difficulties of interpretation and the decision to state that one link is more relevant than another becomes subjective although a scoring system should be objective.

Another scoring scheme was developed which takes account of (i) the hierarchical structure of the concept map, (ii) identified propositions, and (iii) the actual validity versus implied validity of concept map components (Schreiber & Abegg, 1991). This is summarised in the following equality:

\[ X = \left[ x \text{ } n(b + c) \right] + \frac{b}{c} \]  

Where:
- \( X \) is the overall concept map score;
- \( x \) is the initial tally of points awarded for recognition of hierarchical, propositional and valid constructs on a concept map;
- \( n \) is the number of strands in a concept map;
- \( b \) is the summed ratios of number of vocabulary terms to number of hierarchical levels per strand;
- \( c \) is the summed ratios of number of valid connecting lines to total number of connecting lines drawn.
Although Schreiber & Abegg (1991) attempt to address the issue of taking into account the structure of the concept map, they also assume that the practitioner can determine which links are ‘valid’ and which are not.

Equation (i) is very closely related to Equation (ii) in that they are both summations. Aside from a difference in terminology the idea of interconnectedness has been added (Martin et al., 2000). In Equation (i), interconnectedness is represented as the ratio of the score of cross-links to the score of concepts. This is believed to take account of the ‘cohesiveness’ of the concept-map, whereas the score for cross-links perhaps only indicate a measure of ‘integration’ (Martin et al., 2000). However, whatever the nomenclature, the same problem arises of a student who, trained to construct concept maps, may not necessarily understand the concepts contained therein or have ‘ability’ in the topic. Another scoring procedure for scoring structural change (Pearsall, Skipper, & Mintzes, 1997) which bears some similarity to the basic idea behind Kinchin’s method of characterising the structure of concept maps (Kinchin, Hay, & Adams, 2000) is the second system described by (Martin et al., 2000). In Martin et al.’s (2000) method, three structural changes in the knowledge frameworks are documented:

- **Restructuring**, when a concept label is added or deleted from the first hierarchical level of a concept map
- **Accretion**, when ten or more concept labels are added to a pre-existing concept
- **Tuning**, when any change to a pre-existing concept results in a modification of its meaning by the addition of constraining or constant variables.

In order to develop a scoring system, each map is compared to its successor (map 1 v. map 2; map 2 v. map 3; map 3 v. map 4; and so on if appropriate) and a score of 1 or 0 is recorded for each instance or non-instance of a structural change event. At length, each student receives a set of scores. Although important in their own right, these features of a concept mapping assessment system do not answer the problem of how one practitioner may decide certain links and/or cross-links are valid and others not so.

In response to problems such as validity and reliability, **item response theory** (IRT) models were employed for scoring concept maps (Liu, 1994). IRT models have been used for scoring and comparing students’ achievements even if they do not sit the same test (Hambleton, H., & Rogers, 1991). Traditional IRT models have been extended to graded IRT (Samejima, 1969), thus students’ responses to an item no longer have to be scored dichotomously as right or wrong: they can be graded as categories. When IRT is used for scoring concept maps, the four aspects of a concept map, *i.e.*, links, hierarchies, cross-links, and examples (à la Novak & Gowin, 1984), are considered ‘test-items’, and the numbers of links, hierarchies, and examples are considered as students’ categorical responses to the ‘test-items’. By applying graded IRT models to students’ responses, it is possible to obtain students’ ability estimates (Liu, 1994). IRT scoring emphasises the overall structure of students’ concept maps instead of the ‘correctness’ of a specific concept map aspect.

Wilson (1993) analysed the structural characteristics of students’ concept maps using a 24 x 24 matrix representing the inter-relationships between the 24 concepts was produced and non-parametric multidimensional scaling was applied generating co-ordinates on three dimensions. Importantly, the canonical correlation between the co-ordinates and students’ conventional achievement test scores was found to be significant. Liu (1994) claimed that IRT scoring of concept maps is generally valid and reliable and that the correlation between IRT ability estimates and the total concept mapping scores based on Novakian scoring scheme is significant, and thus the Novakian method can be taken to be a reliable indicator of ‘ability’.
Findings

The ‘problem of research’ stated earlier was that concept mapping had a number of inherent problems that researchers tended to overlook such as the validity and reliability of conclusions drawn from the structure / arrangement of the concept maps produced. It is proposed that RGA may move beyond such problems.

What Is the Relationship between RGA and Concept Mapping?

The main aim of this research is to better understand how RGA can be used to study students’ conceptual structures, especially those related to the sciences, and as an example of a concept, PERISSODACTYLA - (odd-toed mammals) was chosen. The technique provides diagrammatic representations of a learner’s conceptual framework in a consistent and rigorous manner. These diagrams can be interpreted as concept maps of a special kind (Lawson, 1997), and the technique allows the concept maps of two or more students to be compared systematically. This is one of the drawbacks of conventional concept mapping where concepts and sub-concepts can be arranged in a hierarchy linearly without labelled linkages (Figure 5) or radially with labelled arrows (Figure 6.) according to preference of the user; however, it is unclear if the user placed any significance on the arrangement, and they were merely being ‘tidy’. If the maps are free drawn, the problem of comparing one child’s concept map with another becomes near insurmountable. Proper comparisons can only be made if the same components are found in all the maps to be compared.

Figure 5: Hierarchical concept map of concept: Perissodactyla drawn using ConceptDrawPro™.
Only 4 out of eighteen possible characters were employed to reduce complexity. The fully drawn concept map is extensive, complicated and difficult to interpret. The lines connecting animals denotes a relationship connecting them, that being the feature within the oval interrupting the link between the animals. The classic concept map generally shows a mere presence/absence relationship, thus a binary state, and usually in the ‘direction’ of what positively contributes to defining that animal. This particular map is in a hierarchical manner with the Order Perissodactyla at the top, Genera in the middle and Species at the bottom. This is an artificial arrangement and the physical distance between the boxes is irrelevant, for example Equus quagga is no more relevant or part of Equus spp. than E. burchelli. Also, the features, in ovals, do not have simple one-to-one connection between Equus spp. and the species at the bottom of the diagram. According to (Novak & Gowin, 1984) the links can be scored to introduce a quantitative element, however this is a sum of the number of links and thus a measure of ‘complexity’, rather than understanding. Repertory grid analysis ‘scores’ all the features against all the constructs and by producing a grid or matrix, mathematical procedures are carried out such as principal component analysis. Atran (1999) applied such a procedure, namely, multidimensional scaling, to an indigenous tribal group’s conception of snakes living near their location: again, the structure is psychological rather than phylogenetic. We are concerned with how such techniques can be used to construct psychological representations of a learner’s conceptual framework but in formal academic biology, principal components analysis is employed to construct representations of the hereditary relationships between living things.
What is the Relationship between a Dendrogram Produced during RGA and Formal Scientifically Derived Taxonomic Trees?

Using the FOCUS program in RepGrid™ (Figure 7.) a dendrogram is produced that has the initial appearance of a taxonomic tree, but is it really what it is? It is a diagram rating the similarity between the animals listed. The distance along the scale to the right hand side is a similarity rating e.g., *E. onager*, *hemionus* and *kiang* are clustered very closely since they are all species of wild ass; and *E. burchelli*, *zebra* and *grevyi* since they are all species of zebra. *E. caballus*, i.e., horse, is rated very close to *Equus spp.* - the generic ‘horse’ – even though all the animals beginning with the abbreviation E. for Equus are equids, i.e., horse-like animals. *Rhinoceros* and *Tapirus* are the next closely related cluster but they are relatively distantly related to each other; and have a greater perceptual, if not psychological, distance from horses, zebra and the like. Odd ones out are the extinct guagga, an animal which belongs close to zebra, and *E. asinus*, i.e., donkey, though this last species is perceptually closest to the wild asses.

Using the same data, principal components analysis can be done to extract loadings which can be plotted on a Cartesian plane (Figure 8.). In RGA, the analysis is done twice so that the elements are plotted as points and the constructs as vectors. The distance of the element from the construct vector, which is polar, is related to the relevance that the participant thinks about this context, as was inferred earlier in the introduction to RGA.
Kitching et al. (1998) believe that any comparative data, and cite linguistics as an example, can be organised using cladistics, the end result of which is a cladogram (Figure 9). They define cladistics as a method of classification that groups taxa hierarchically into discrete set and subsets. There is an obvious relationship between arranging the concepts of living things using a repertory grid as the raw data and devising an arrangement of concepts of living things using biochemical or morphological information as the raw data. In the former, the data is personal or mental. In the latter case the data is physical. It is interesting to note however that the same system can be used to elicit organisational frameworks of biological concepts. It is both important and interesting that both repertory grid analysis - employing personally constructed mental data to produce graphs resembling concept maps; and cladistics - employing derived physical data to produce branching diagrams called cladograms; both share a similar underlying principle. The relationship between concepts is relative. This does alter the trend in dealing with scientific concepts whether physically or psychologically described, in that they have tended in the past to be seen as absolutes, even though the relationship between specific concepts remained vague.
The comparison of RGA matrices have been dealt with elsewhere (McCloughlin & Matthews, 2005, 2010), and here, suffice to say that since matrices may be treated as single entities in mathematics, the simplest way to compare two matrices is to subtract them and observe the results of subtractions. Two matrices could involve an expert and a novice to locate areas of improvement, and repeated to note development. There are also a number of ways to compare large numbers of matrices of 30, 40 or more individuals; one, is to average the matrices and analyze the resultant ‘averaged’ matrix; averaging may be simple averaging of the matrices but more often a centering technique is used.

Conclusions

The principal components plot of RGA is a special sort of concept map, but more so since the distance between concepts has a mathematical significance. The principal components analysis plot locates constructs and their elements, thus rather than a simple propositional logic evident in simple concept maps, a multirelational logic inbuilt in the mathematics of principal components analysis becomes evident in principal components analysis plots. RGA provides for meaningful assessment rather than scoring whole concept maps ‘in the round’, a so-called ‘holistic’ score, or taking the four main components and scoring for ‘relevance’, by which is meant the ‘nearest to the educators idea’ as opposed to the relevance the student places on the concept.

Hierarchies and complexities may be drawn into concept maps that were not
intended or failure to grasp the ‘rules’ of the concept map rubric would result in biased scoring. RGA bypasses these issues, and a proper ‘holistic’ score is achieved by RGA, not in the form of a single integer but rather as a matrix. Thus the problem of quantification of concept mapping is addressed. The meaningful comparison of concept maps between students, or between students and educators is possible (whereas this has proven to be one of the more problematic areas in ‘classical’ concept mapping). Although, a number of positive things can be said about RGA, the one major drawback is that the principle components plots need to be interpreted by a person who has spent some time getting to know the nuances of the plots. However, it can be done relatively easy, although it is not ‘intuitive’ unlike may things to do with computers.

The revised Irish curriculum for primary schooling may state that concept mapping is an appropriate form of assessment in science but it does not give guidance on how concept mapping is to be done, or indeed what sort should be employed. Certainly, much more research needs to be done in repertory grid analysis and its implications and applications in curricular research have yet to be fully explored.

References


Kubrusly, M. H. (1984). *Does the teacher make a difference? A study on the correlation between learners’ perceptions of their teachers and their reactions to the topics covered during the course.* (Master of Arts), Lancaster.


Nespor, J. K. (1985). *The role of beliefs in the practice of teaching: final report of the Teacher Beliefs Study*. Austin, Texas: Research and Development Center for Teacher Education, Texas University at Austin.


