Exploring the Mobile Structural Assessment Tool: Concept Maps for Learning Website

Exploración de la herramienta de aseguramiento estructural móvil: mapas conceptuales para websites de aprendizaje

Mehmet Filiz[1,a], David Trumpower[1,b], Arun Vanapalli[1,c]

[1]Measurement Evaluation and Assessment Research Unit, Faculty of Education, University of Ottawa, Canada,

———————————————

Abstract

In this paper, we describe how the pathfinder algorithm converts relatedness ratings of concept pairs to concept maps; we also present how this algorithm has been used to develop the Concept Maps for Learning website (`www.conceptmapsforlearning.com`) based on the principles of effective formative assessment. The pathfinder networks, one of the network representation tools, claim to help more students memorize and recall the relations between concepts than spatial representation tools (such as Multi-Dimensional Scaling). Therefore, the pathfinder networks have been used in various studies on knowledge structures, including identifying students' misconceptions. To accomplish this, each student's knowledge map and the expert knowledge map are compared via the pathfinder software, and the differences between these maps are highlighted. After misconceptions are identified, the pathfinder software fails to provide any feedback on these misconceptions. To overcome this weakness, we have been developing a mobile-based concept mapping tool providing visual, textual and remedial feedback (ex. videos, website links and applets) on the concept relations. This information is then placed on the expert concept map, but not on the student's concept map. Additionally, students are asked to note what they understand from given feedback, and given the opportunity to revise their knowledge maps after receiving various types of feedback.

Key words: Concept Maps, Effective Feedback, Pathfinder Network, Structural Assessment.

———————————

[a]PhD. Student. E-mail: mehmetfiliz52@hotmail.com
[b]Associate Professor. E-mail: david.trumpower@uOttawa.ca
[c]M.A. Student. E-mail: vanapalliarun@gmail.com

Resumen

En este artículo se describe cómo el algoritmo de búsqueda de ruta convierte puntajes de conceptos pareados en mapas conceptuales. También se presenta cómo este algoritmo ha sido utilizado para desarrollar estos mapas conceptuales para aprendizaje (`www.conceptmapsforlearning.com`) basados en los principios del aseguramiento formativo efectivo.

Las redes de búsqueda de ruta, una de las herramientas de representación de redes, ayudan a memorizar a los estudiantes y enunciar las relaciones entre mapas más que las herramientas de expresión espacial (tales como el escalonamiento multidimensional). Por tanto, las redes de búsqueda de rutas han sido usadas en varios estudios de estructura del conocimiento incluyendo la identificación de malos conceptos usados por los estudiantes. Para lograr esto, cada mapa de conocimiento tanto del estudiante como del experto son comparados vía el software de búsqueda de ruta y se remarcan las diferencias entre éstos. Después que los malos conceptos son identificados, el software de búsqueda falla en entregar una retroalimentación en estos nodos conceptuales. Para superar esta debilidad, se desarrolla una herramienta de mapa conceptual móvil que manda retroalimentaciones visuales, textuales y remediales (e.g. vídeos, enlaces a páginas web y applets) en las relaciones de los conceptos. Adicionalmente, los estudiantes son preguntados acerca de qué entienden de la retroalimentación brindada y se les da la oportunidad de revisar sus mapas de conocimiento después de recibir varios tipos de retroalimentación.

Palabras clave: aseguramiento estructural, mapas conceptuales, redes de bśqueda de ruta, retroalimentación efectiva.

1. Introduction

The most important factor in learning is to identify learners' existing knowledge. Meaningful learning relies on how well connections are made between what is known and what is still to be known (Ausubel 1978). This prior knowledge can be presented either spatially or non-spatially via digital knowledge mapping tools (Figure 1).

In spatial models (e.g., Multi-Dimensional Scaling (MDS)), concepts are placed (based on distance) in a geometric space of continuous attributes or dimensions. In contrast, concepts are located based on the presence or absence of shared discrete attributes in non-spatial models (e.g., Pathfinder Network Representation) (Cooke 1992b). As compared to MDS, Pathfinder Network Representation has several advantages. First, concepts do not need to be placed hierarchically in Pathfinder Network Representation (Cooke 1992b). Second, asymmetrical relationships among concepts can be visualized via Pathfinder Network Representation (Cooke 1992b). Third, Pathfinder Network Representation better captures information related to recall than MDS (Cooke, Durso & Schvaneveldt 1986). Finally, and most importantly, while MDS illustrates global relations among concepts, Pathfinder Network Presentation demonstrates local relations among concepts (Cooke & Schvaneveldt 1988, Chen 2004). Due to these strengths, Pathfinder
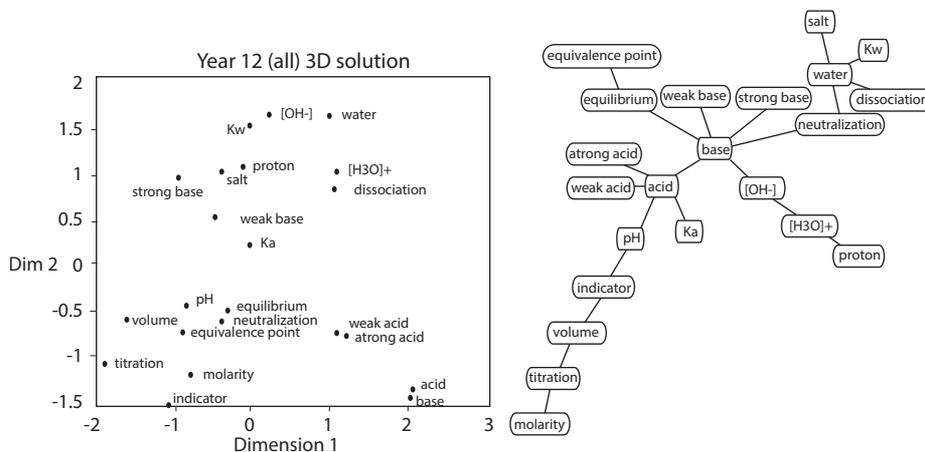
FIGURE 1: Generated spatial and non-spatial representation of the same data (Wilson, 1998).

Network Representation may provide a better approach to identifying students' specific incomplete understandings, including their misconceptions, than MDS. Therefore, Pathfinder Network Representation was chosen as the basis for visualizing students' knowledge in the mobile structural assessment tool discussed in this paper.

2. Pathfinder Software and Structural Assessment

Pathfinder software generates network representations, which consist of nodes and lines, based on obtained proximity data. Networks generated by Pathfinder are similar to concept maps without linking terms; nodes and lines in the networks correspond to concepts and relationships, respectively. Therefore, networks produced by Pathfinder will be referred to as knowledge maps.

Structural assessment using Pathfinder involves three basic steps: obtaining proximity data from a student, converting these data into a knowledge map, and comparing the student's knowledge map with an expert knowledge map (Goldsmith, Johnson & Acton 1991, Taricani & Clariana 2006, Kim 2012). In this regard, we first reveal different ways of collecting proximity data. Then, how these data are converted to a knowledge map via the Pathfinder Scaling Algorithm is illustrated. Thereafter, criteria for generating ideal expert knowledge maps are identified. Finally, a brief literature review is given for demonstrating the ways that students' knowledge maps are evaluated using expert knowledge maps.

2.1. Obtaining Proximity Data

Two approaches are used for collecting proximity data to create students' knowledge maps via the Pathfinder Scaling Algorithm. In the first approach, stu-

dents are given a set of concepts, and then asked to consider concept associations for sorting the concepts, making an ordered recall or rating the degree of relatedness of concept pairs. Among these tasks, performing pairwise ratings has been shown to be superior to the alternative tasks (Rowe, Cooke, Hall & Halgren 1996).

In the second approach to obtaining proximity data, students are asked to prepare a document (e.g., write an essay) without a predefined set of concepts (Chen 2011, Cooke, Neville & Rowe 1996, Davis, Curtis & Tschetter 2003, DeChurch & Mesmer-Magnus 2010, Goldsmith et al. 1991). Proximity data can then be computed from the document using the Analysis of Lexical Aggregates (ALA-Reader) software (Koul, Clariana & Salehi 2005). ALA-Reader requires teachers to create a set of concepts (maximum of 30 words) consisting of important terms and their synonyms and metonyms. Afterwards, this concept set and students' essays are submitted to the software for converting the co-occurrences of terms into propositions. Finally, these propositions are compiled across all sentences into a proximity array. Studies have shown that the grades given by ALA-Reader are moderately consistent with those given by human evaluators (Clariana & Koul 2004, Koul et al. 2005, Clariana & Wallace 2007). On the other hand, this technique has some limitations. First, Kim (2013) points out that ALA-Reader cannot extract information about why two linked concepts are related. In addition, Gomez, Hadfield & Housner (1996) argue that the pairwise ratings task is a more direct approach to capturing knowledge structures than essay writing, because students are required to think about every possible concept relationship while performing pairwise ratings. Consistent with Gomez et al. (1996)'s argument, Boring (2005) found that knowledge structures of students in an introductory physiology course derived from a pairwise rating task were more coherent than when derived from essays.

2.2. How the Pathfinder Algorithm Works

The Pathfinder Scaling Algorithm determines whether two concepts are linked through the triangle inequality (Cooke 1992b, Guerrero-Bote, Zapico-Alonso, Esinosa-Calvo, Gomez-Crisostomo & Moya-Anegon 2006, Nash & Nash 2003). This algorithm has two parameters. The parameter r is related to the Minkowski metric which is used for calculating indirect path distances between two concepts, and its value can be between 1 and infinity (Kivlighan & Tibbits 2012, Nash & Nash 2003, Schvaneveldt 1990). While the city block metric is computed when r is equal to 1, the length of an indirect path equals the maximum weight of the links that create the path when r is set as infinity, (Cooke 1992b, Nash & Nash 2003). The other parameter, q, determines the maximum number of links in examining any path, and this parameter could be set from 1 to the number of nodes minus one (Cooke 1992b, Kivlighan & Tibbits 2012, Nash & Nash 2003, Schvaneveldt 1990). After both parameters are set, two concepts are assigned to be linked if the length of any indirect path is equal to or longer than the length of the link between these two concepts (Nash & Nash 2003).

When the value of r or q is increased, the number of the links in a network might be decreased (Chen 1998, Nash & Nash 2003). Therefore, r and q set to

infinity and *n*-1 (*n* referring to the number of concepts), respectively, will create the most parsimonious network having the least number of links (Chen 1998, Cooke 1992*b*, Guerrero-Bote, Zapico-Alonso, Espinosa-Calvo, Gomez-Crisostomo & Moya-Anegon 2006, Kivlighan & Tibbits 2012, Nash & Nash 2003). An example of how the Pathfinder algorithm works under such conditions is demonstrated next.

Pathfinder creates a knowledge map from pairwise ratings in six steps. These steps are illustrated via the following example ratings (see Table 1) from Trumpower & Sarwar (2010).

TABLE 1: An example of pairwise ratings.

	Less Related			More Related	
Milk - Barn	1	2	3	**4**	5
Barn - Cow	1	2	3	**4**	5
Cow - Tractor	**1**	2	3	4	5
Tractor - Milk	**1**	2	3	4	5
Cow - Milk	1	2	3	4	**5**
Tractor - Barn	1	2	**3**	4	5

1. A matrix (see Table 2) is created by using the ratings above.

TABLE 2: Created matrix based on the pairwise ratings.

	Ratings table				Generated matrix			
	Milk	Barn	Cow	Tractor				
Milk	0	4	5	1	$\begin{bmatrix} 0 & 4 & 5 & 1 \\ 4 & 0 & 4 & 3 \\ 5 & 4 & 0 & 1 \\ 1 & 3 & 1 & 0 \end{bmatrix}$			
Barn	4	0	4	3				
Cow	5	4	0	1				
Tractor	1	3	1	0				

2. The ratings in the matrix above are recoded because there is a reverse relationship between the relatedness degree of concept pairs and the distance between them (1→5, 2→4, 3→3, 4→2, 5→1).

$$\boldsymbol{W} = \begin{bmatrix} 0 & 2 & 1 & 5 \\ 2 & 0 & 2 & 3 \\ 1 & 2 & 0 & 5 \\ 5 & 3 & 5 & 0 \end{bmatrix}$$

3. $\boldsymbol{W}^3$ is created. In order to do this, first $\boldsymbol{W}^2$ must be calculated. The equation (1) is used for calculating each element of $\boldsymbol{W}^2$.

$$w_{jk}^q = ((w_{jm})^r + (w_{mk}^{q-1})^r)^{1/r} \tag{1}$$

For $r=\infty$, the result of the equation above is equal to $\max(w_{jm}, w_{mk}^{q-1})$ (Ichino & Yaguchi 1994). In this equation, *j*, *k*, and *m* are the indexes representing the order of concepts in the set. For instance, *j* and *k* equal to 1 and 2,

respectively, when that equation is for computing value of the element for Milk and Barn concepts. In addition, m can be 3 or 4 because there are four concepts in the set and it cannot be equal to j and k. Thus, $\boldsymbol{W}^2_{12}$ is computed as shown below.

$r \to \infty \; w^2_{12} = ((w_{13})^r + (w_{32})^r)^{1/r} = ((1)^r + (2)^r)^{1/r} = 2$

$r \to \infty \; w^2_{12} = ((w_{14})^r + (w_{42})^r)^{1/r} = ((5)^r + (3)^r)^{1/r} = 5$

Therefore, $\boldsymbol{W}^2_{12}$ is equal to 2, because it is equal to minimum value of $\boldsymbol{W}^2_{12}$'s.

$$\boldsymbol{W}^2 = \begin{bmatrix} 0 & 2 & 2 & 3 \\ 2 & 0 & 2 & 5 \\ 2 & 2 & 0 & 3 \\ 3 & 5 & 3 & 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{W}^3 = \begin{bmatrix} 0 & 2 & 2 & 3 \\ 2 & 0 & 2 & 3 \\ 2 & 2 & 0 & 3 \\ 3 & 3 & 3 & 0 \end{bmatrix}$$

4. $\boldsymbol{D}^3$ is computed by taking the minimum value for each element from $\boldsymbol{W}$ and $\boldsymbol{W}^3$.

$$\boldsymbol{W} = \begin{bmatrix} 0 & 2 & 1 & 5 \\ 2 & 0 & 2 & 3 \\ 1 & 2 & 0 & 5 \\ 5 & 3 & 5 & 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{W}^3 = \begin{bmatrix} 0 & 2 & 2 & 3 \\ 2 & 0 & 2 & 3 \\ 2 & 2 & 0 & 3 \\ 3 & 3 & 3 & 0 \end{bmatrix} \to \boldsymbol{D}^3 = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 2 & 0 & 2 & 3 \\ 1 & 2 & 0 & 3 \\ 3 & 3 & 3 & 0 \end{bmatrix}$$

5. Each element from $\boldsymbol{D}^3$ and $\boldsymbol{W}$ is compared to compute matrix $\boldsymbol{L}$. If two compared elements are equal, the value of the corresponding element in matrix $\boldsymbol{L}$ is "1" which indicates a link between two concepts. If two compared elements are not equal, the value of the corresponding element in matrix $\boldsymbol{L}$ is "0" which reveals that two concepts are not directly related.

$$\boldsymbol{W} = \begin{bmatrix} 0 & 2 & 1 & 5 \\ 2 & 0 & 2 & 3 \\ 1 & 2 & 0 & 5 \\ 5 & 3 & 5 & 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{D}^3 = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 2 & 0 & 2 & 3 \\ 1 & 2 & 0 & 3 \\ 3 & 3 & 3 & 0 \end{bmatrix} \to \boldsymbol{L} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

6. According to matrix $\boldsymbol{L}$ above, Milk is linked to Barn, Milk is linked to Cow, Barn is linked to Cow, and Barn is linked to Tractor.

This referent knowledge map can also be compared with another knowledge map as demonstrated below:

1. Assume that the ratings below (see Table 3) were generated by a student.

TABLE 3: Student's pairwise ratings.

	Milk	Barn	Cow	Tractor
Milk	0	1	5	4
Barn	1	0	4	3
Cow	5	4	0	1
Tractor	4	3	1	0

2. Based on the ratings above, matrix $\boldsymbol{L}$ and then the student's knowledge map are created (Figure 2).
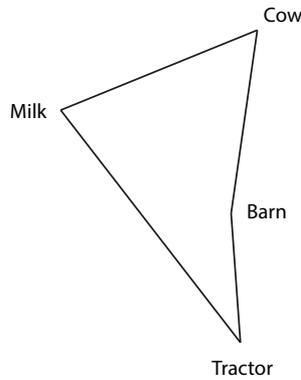
FIGURE 2: The student's knowledge map.

3. Matrix $\boldsymbol{L}$ of the referent knowledge map (L_e) and matrix $\boldsymbol{L}$ of the student's knowledge map (L_s) are compared. Then, the similarities and differences are identified as illustrated in the Table 4:

TABLE 4: Criteria for generating links in the visual feedback map.

Referent knowledge map (L$_e$)	Student knowledge map (L$_s$)	The feedback knowledge map (L$_f$)
A link does not exist (0)	A link does not exist (0)	A link does not exist (0)
A link exists (1)	A link exists (1)	A relevant link (1)
A link does not exist (0)	A link exists (1)	An extraneous link (2)
A link exists (1)	A link does not exist (0)	A missing link (3)

4. The Matrix $\boldsymbol{L}$ (see Table 5) of the feedback knowledge map $(\boldsymbol{L}_f)$ matrix is created as described earlier.

TABLE 5: The Matrix $\boldsymbol{L}$ of the feedback knowledge map.

The referent knowledge map	The student knowledge map	The feedback knowledge map
$\boldsymbol{L}_e = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$	$\boldsymbol{L}_s = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$	$\boldsymbol{L}_f = \begin{bmatrix} 0 & 3 & 1 & 2 \\ 3 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{bmatrix}$

5. Finally, based on the matrix $\boldsymbol{L}_f$, the feedback map is generated (Figure 3).
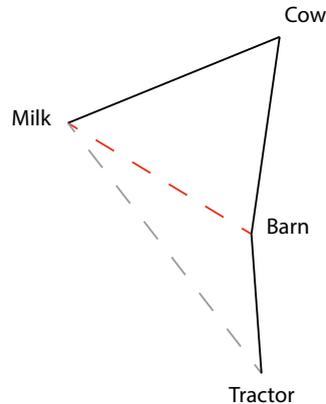

FIGURE 3: An example of response format.

2.3. Criteria for Generating Referent Knowledge Maps

Creating referent knowledge maps, which best reflect the structure of a particular topic, is one of the essential components of the Pathfinder structural assessment technique, because these maps are used for evaluating students' knowledge maps (Acton, Johnson & Goldsmith 1994, Cooke 1999). Therefore, a few criteria should be set for generating valid referent knowledge maps, including limiting the number of concepts in the concept list, and testing the accuracy of expert ratings.

First, the number of concepts in a knowledge map should be limited even though Goldsmith et al. (1991) claim that the greater the number of concepts in a set, the more accurate the knowledge network. Goldsmith et al. (1991) used a set of 30 concepts in their study, and each student spent approximately one hour to complete the pairwise ratings task. Consequently, having more concepts in a knowledge map results in time-consuming pairwise ratings tasks, and this might draw students' attention away from said task (Dicerbo 2007). Trumpower & Sarwar (2010), therefore, suggest that the set of concepts should consist of at most around 20 concepts, whereas Casas-García & Luengo-González (2013) claim that there should be no more than 12 concepts in the set. On the other hand, Cooke (1999) and Goldsmith et al. (1991) argue that a set must include at least 12 concepts to accurately assess a domain. Thus, it appears that a set of concepts should consist of around 12 concepts to provide an accurate measure of a domain. However, a few studies have shown that less than 12 concepts might be enough to accurately assess a topic if there are core concepts confirmed by domain experts. For instance, Trumpower & Sarwar (2010) used 11 concepts of a unit on work, energy, and power to capture students' conceptual understanding before and after a feedback/remediation activity. After 24 high school students rated the relatedness degree of these concept pairs before and after the feedback/remediation activity, their pre and post knowledge maps were generated. Then, the similarity indexes

for their pre and post knowledge maps were calculated by comparing them with the expert knowledge map. The mean of the similarity values rose from 0.45 (SD = 0.11) to 0.56 (SD = 0.12), and this increase was significant ($F(1, 23) = 17.04$; $p < 0.0001$). Likewise, Casas-García & Luengo-González (2013) created a set of 11 concepts for generating students' knowledge maps on the mathematical concept of angles. Primary (Grade 3P to 6P), Secondary (Grade 1E to 4E), Olympiad (2E), and Mathematics Undergraduate students were asked to perform the pairwise ratings of these concepts. Then, the representative knowledge map of each group was created and compared to each other for computing similarity indexes between the knowledge maps. Thereafter, a Kruskal-Wallis test was performed, and it was found that the similarity between the knowledge maps increased with rising age and experience of the students ($H = 15.252$, 7 d.f., $p = 0.0329$). Additionally, Boring (2005) identified 10 core concepts of an introductory psychology course to compare students' essay scores (g), which are given by the human graders (holistic and analytic) and the computerized grader, with Pathfinder similarity index (C). For the holistic graders, $g = 0.414C$ - 0.048, where $R^2 = 0.001$ and $p > 0.05$. Therefore, the degree of Pathfinder network similarity was not a good indicator of the grade awarded by the holistic graders. Similarly, it was found that the relationship between the Pathfinder network similarities and the score awarded by the analytic graders was not significant ($R^2 = 0.012$ and $p > 0.05$). On the other hand, the degree of similarity between the computerized grader's Pathfinder network and the student essay writer's network was a reasonably good indicator of the grade awarded for that paper ($R^2 = 0.247$ and $p = 0.06$). In these studies, although the researchers used less than 12 concepts, they assumed that generated knowledge maps were valid measures of students' understanding. Furthermore, at least three experts should be asked to rate the relatedness degree of concept pairs; then, the mean of their ratings should be calculated for each concept pair to create more valid expert knowledge maps (Wilson 1998, Day, Arthur Jr & Gettman 2001, Sarwar 2012, Kivlighan & Tibbits 2012). However, some studies reveal that expert ratings may not be consistently accurate for several reasons. These reasons include having different types of domain expertise or teaching at different levels. For instance, Moni, Beswick & Moni (2005) asked two experts (an experienced physiologist and a pharmacologist) to score concept maps related to physiology, and found that there were larger differences between their scores for some of these concept maps. Moreover, Von Minden, Walls & Nardi (1998) concluded that the knowledge maps of elementary school mathematics teachers showed strongest agreement with high school teachers and weakest agreement with university mathematicians. Thus, not all expertise is equal. When creating a referent knowledge map from the mean ratings of several experts, such inequalities must be taken into account. Sarwar (2012) provides an example of how to deal with differences and inaccuracies in expert ratings by determining the reliability within each expert rating and the reliability across experts' ratings. For the reliability within each expert's set of ratings (i.e., test-retest reliability), experts were asked to rate the relatedness degree of all concept pairs and, in addition, the first five concept pairs were repeated at the end of the survey. For example, if there are six concepts in a set, the number of possible concept pairs is 15. In addition to these

pairs, the first five pairs are repeated so there are 20 concept pairs (see Figure 4). Thus, the test-retest reliability of the ratings of the first five and last five pairs for each expert can be computed.

First Concept	Second Concept	1	2	3	4	5
F-ratio	Number of groups	0	0	0	0	0
Between-groups mean square	Sum of squares between groups	0	0	0	0	0
Within-groups mean square	Degree of freedom	0	0	0	0	0
Number of groups	Degree of freedom	0	0	0	0	0
Sum of squares between groups	Degree of freedom	0	0	0	0	0
Between-groups mean square	Within-groups mean square	0	0	0	0	0
F-ratio	Within-groups mean square	0	0	0	0	0
Between-groups mean square	Degree of freedom	0	0	0	0	0
Number of groups	Between-groups mean square	0	0	0	0	0
Number of groups	Sum of squares between groups	0	0	0	0	0
F-ratio	Sum of squares between groups	0	0	0	0	0
F-ratio	Between-groups mean square	0	0	0	0	0
Within-groups mean square	Sum of squares between groups	0	0	0	0	0
F-ratio	Degree of freedom	0	0	0	0	0
Number of groups	Within-groups mean square	0	0	0	0	0
Number of groups	F-ratio	0	0	0	0	0
Between-groups mean square	F-ratio	0	0	0	0	0
Degree of freedom	F-ratio	0	0	0	0	0
Sum of squares between groups	F-ratio	0	0	0	0	0
Within-groups mean square	F-ratio	0	0	0	0	0

FIGURE 4: An example of response format.

As well, Pearson correlation coefficients were computed between each pair of experts' set of ratings to measure how consistent the experts were (i.e., inter-rater reliability). Following the calculation of these two measures of reliability, the author eliminated any expert's ratings that did not show acceptable test-retest reliability or that did not correlate well with the other experts' ratings. A cut-off value of .50 was used in both cases. After these two reliability checks, the average value of the remaining experts' ratings for each concept pair was used to generate the referent knowledge map. In summary, reliability within each expert's ratings as well as high correlation across expert ratings is crucial for generating accurate expert knowledge maps. In conclusion, limiting the number of concepts in the concept list and testing the reliability of expert ratings are basic criteria for generating ideal expert knowledge maps.

2.4. Brief Review of Studies on Pathfinder Networks

When a student's knowledge map is compared with an expert knowledge map through Pathfinder software, three different measures (PRX, GTD and C) are computed. PRX demonstrates the correlation of raw proximities whereas GTD

(graph-theoretic distance) illustrates the correlation of the distances between the nodes in two knowledge maps (Lau & Yuen 2010). C (configural similarity) is computed by dividing the number of common links in two knowledge maps by the total number of links in both knowledge maps (Clariana 2010). Among these measures, C is most commonly used.

McGaghie, McCrimmon, Mitchell, Thompson & Ravitch (2000) examined the correlation between the configural similarities of students' knowledge maps with their examination performances. In this study, 153 Northwestern University medical students and 76 University of Wisconsin veterinary medical students were asked to perform pairwise ratings of 12 concepts before a three-week-long instructional unit on pulmonary physiology. After this instruction, students took an examination and rated the relatedness degree of the same 12 concepts again. Then, the configural similarities of students' pre-knowledge maps and post-knowledge maps with an expert knowledge map were computed. They found that students' knowledge maps became more similar to the expert knowledge map following instruction, increasing from .07 to .22 for the Northwestern University students and from .12 to .18 for the University of Wisconsin students. The authors also noted, however, that the correlation between the configural similarity of students' post-knowledge maps with the expert knowledge map, and their examination performance, was not statistically significant at the $p = .05$ level. The reason behind this unexpected result may be the difference between the type of knowledge measured by configural similarity and the examination (see, for example, Taricani & Clariana 2006).

Comparison of students' knowledge maps with an expert knowledge map based on the presence and absence of concept relations may provide further information about students' knowledge, such as their incomplete understanding and/or misconceptions. Dicerbo (2007) compared a student map (utilizing the mean of several students' results) with an expert knowledge map and identified how experts' knowledge organization differs from students' knowledge organization. In this study, 156 students and 12 teachers were asked to perform a relatedness task of two sets of concepts prior to instruction. One set represented the course as a whole, whereas the other consisted of concepts from a specific chapter (*Networking Fundamentals*) within a course. These ratings were then submitted to the Pathfinder software to generate the knowledge maps of the students and experts as well as to find the C index between them. While the C index between students' knowledge map and experts' knowledge map was .27 for the set of concepts representing the entire course, the C index between students' knowledge map and experts' knowledge map was .23 for the concepts of *Networking Fundamentals*. Furthermore, the expert map for the concept set representing the course as a whole corresponded to the Open System Interconnections model, whereas the tendency of students was to categorize concepts as equipment and words that sounded similar. Additionally, while experts organized the concepts of *Networking Fundamentals* by network type (local "*LAN*" or wide-area "*WAN*"), students categorized these concepts in a device-centric organization (Figure 5). Even though this study might give some clues to teachers for targeting their instruction, it does not provide any information regarding *individual* students' understandings of particular concept relations.
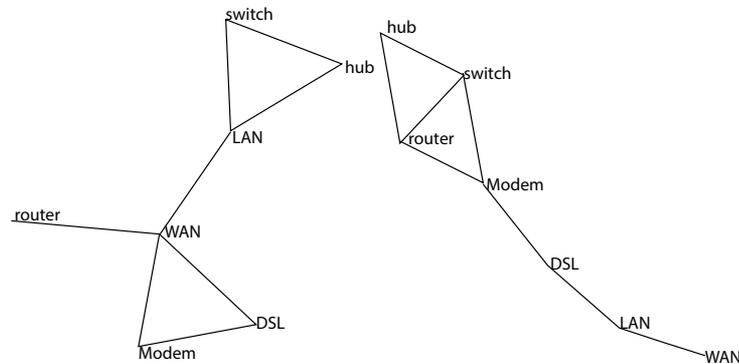
FIGURE 5: Expert's knowledge map and Student's knowledge map for the concepts of
Networking Fundamentals.

Similarly, Boldt (2001) compared a knowledge map of introductory financial
accounting students and a knowledge map of Masters-level students with an ex-
pert knowledge map. First, Pearson product-moment correlation coefficients were
calculated for each student's ratings with every other student's ratings. Then, the
mean coefficients of relatedness ratings between students group were computed.
It was found that the knowledge map of the Masters-level students was more sim-
ilar to the expert knowledge map (.48) as compared to the knowledge map of the
introductory financial accounting students (.20). In addition, the concepts in the
knowledge map of the Masters students and introductory financial accounting stu-
dents were classified into four categories as *well-defined, misdefined, overdefined
and underdefined*. A *well-defined* concept refers to any concept which is linked to
the same concepts in both the student knowledge map and the expert knowledge
map. The concept is *misdefined* if there are no common links between the stu-
dent knowledge maps and the expert knowledge map, yet it has extraneous links
which are not present in the expert knowledge map. *Overdefined* concepts have
both all present links in the expert knowledge map and some extraneous links. In
contrast, *underdefined* concepts in a student knowledge map only have some of
the links which are present in the expert knowledge map. Boldt (2001) found that
the knowledge map of the Master's students had more well-defined concepts than
did the knowledge map of the introductory financial accounting students. The
results of this study informed the course instructors to revise their lesson plans for
the next courses. Again, however, *individual* students, *well-defined, misdefined,
overdefined and underdefined* concepts were not diagnosed.

Kivlighan & Tibbits (2012) went one step further, identifying differences be-
tween each counseling trainee's knowledge structure and an expert knowledge
structure in order to investigate trainees' misconceptions with respect to coun-
seling. Four experienced group therapists and fifty group counseling trainees were
asked to use group counseling interventions for addressing a variety of group coun-
seling situations adapted from the Group Therapy Questionnaire. Data obtained
from the interventions were converted to knowledge maps via Pathfinder software.
First, each trainee's knowledge map was compared to the expert counselors' knowl-
edge map, and C index was computed which was ranged from .04 to .31 ($M = .15$,
SD $= .06$). Then, the discrepancies between the trainees' knowledge structures

and the expert knowledge structure were classified into two categories, as errors of omission or errors of commission. The link is called an error of commission if two interventions are linked in trainees' knowledge structures, but not in the expert's knowledge structure. On the other hand, the link is identified as an error of omission provided that the link between two interventions is only present in the expert's knowledge structure. More than 75% of trainees had nine repeated errors of omission, and only one frequently appearing error of commission (silence to attack). As a result of cluster analysis, four distinct subgroups within the group of trainees were identified. This finding might help trainers better focus their instructional activities to trainees' misconceptions.

These studies show that Pathfinder software can be used for identifying students' incomplete understandings in four steps. First of all, a set of concepts to be assessed must be identified. Then, an expert structure should be generated based on several experts' ratings. The next step is to obtain students' structures. Finally, each student's knowledge structure is compared with the expert knowledge map to find differences between the two knowledge structures. Although Pathfinder software identifies students' incomplete understandings, it fails to help students to act on their partial understandings because of several limitations. For example, Cooke (1992a) reveals that Pathfinder software does not allow users to label links. In addition, even though the expert and student structures share the same linked concept pairs, it does not mean that a student's understanding is at the expert level. Dicerbo (2007) also indicates this limitation and calls it a tradeoff of the software that provides quantitative results to teachers about their students' knowledge. Thus, identifying students' misconceptions via Pathfinder software is possible, but is not done automatically by the software.

In another study, after each student's misconceptions were identified, Trumpower & Sarwar (2010) gave specific feedback and/or instructions to help students gain a better understanding of any missing concept relations. Students were first shown both their knowledge structure and an expert knowledge structure, and then asked to reflect on similarities and differences. Additionally, students were required to either solve problems or review examples which are aimed to reduce the number of missing concept relations in their particular knowledge maps. Although effective, the process was time and labor intensive due to the fact that each student's knowledge structure was manually evaluated by the researchers in order to determine the individualized remedial activities (i.e., problems to solve and examples to review) corresponding with that student's missing links and misconceptions. Therefore, the time gap between identifying misconceptions and providing remedial activities may have attenuated the effectiveness of the given feedback.

To overcome these constraints, software programmers have developed some applications based on the Pathfinder network generation algorithm. One such application is our Concept Maps for Learning Website (CMFL). Within the CMFL website, students create knowledge maps on a particular subject and then receive individualized feedback and associated instructional material (e.g., videos, website links, examples, problems, etc.) based on a comparison of their knowledge map and a subject matter expert's map. After studying the feedback and instructional

material, teachers can track their students' progress by having them create revised knowledge maps.

There are four different user identities on the website: administrator, experts, teachers and students. The role of the administrator is to gather sets of concepts on a variety of topics and to create a list of experienced experts on each particular topic. Then, experts are asked to rate the relatedness degree of concept pairs for their topic of expertise. Next, the administrator creates expert knowledge maps based on the mean expert ratings. Thereafter, the expert knowledge maps become available for teachers and they can choose from amongst the topics. Afterwards, students are provided with the concepts corresponding to the chosen topic and rate the degree of relatedness between the concepts in order to generate their knowledge map. The website then compares each student's knowledge map with the expert knowledge map to generate individualized feedback for each student. Feedback is comprised of a visual presentation of the expert knowledge map superimposed over the student's map with any discrepancies highlighted by different types of links. After reviewing this feedback, students are asked to reflect on each missing link. That is, they are asked to think of ways in which the concepts connected by the missing links might be related. Next, students are asked to study textual explanations and linked material to better understand any missing links. After again reflecting on ways in which the concepts connected by missing links are related, students are asked to perform the concept relatedness rating task for a second time in order to generate a revised knowledge map. This new knowledge map can then be compared to the expert knowledge map to see any improved understanding, after which the formative cycle may be repeated as necessary.

In previous studies, we have described how the CMFL tool conforms with established principles of assessment for learning, how it could be used in statistics education, how it contributes to assessment for learning practices, and how it differs from other comprehensive digital knowledge map-based formative assessment systems (Filiz, Trumpower & Atas 2012, Filiz, Trumpower & Atas 2013b, Filiz, Trumpower & Atas 2013a, Trumpower, Filiz & Sarwar 2014).

The personalized diagnosis and remedial learning system (PDRLS) (Chen 2011) is another tool developed to capture students' knowledge structure, identify students' misconceptions by comparing their knowledge structure with an expert structure, and provide individual feedback relying on remedial learning paths much like our own CMFL website. Chen (2011) conducted an experimental study to determine the quality and effectiveness of PDRLS on improving students' learning performance. Students who were enrolled in a JAVA programming design course, were randomly grouped into an experimental group ($n = 72$) and control group ($n = 73$). After the assigned topic was taught, students created their knowledge map via PDRLS and answered a pre-test consisting of 10 questions. Then, students in the experimental group were asked to study remedial feedback for 30 minutes, whereas those in the control group only received their score based on their knowledge structure. Finally, all participants took a ten-question multiple-choice post-test. The results of this study indicated that learners in the experimental group achieved more improved learning performance and self-efficacy than those in the control group, and students with a lower level of knowledge received a

greater benefit from the PDRLS remedial feedback than those with a higher level of knowledge.

Although very similar, the PDRLS and our CMFL website have a few notable differences. For example, students are shown their knowledge map and the expert knowledge map on the same page in the PDRLS. In CMFL, however, students receive individualized visual feedback which is comprised of a visual presentation of the expert knowledge map superimposed over the student's, with any discrepancies highlighted by different types of links. Moreover, students receive instructional materials through the diagnosed remedial learning path in the PDRLS. On the other hand, in CMFL, instructional feedback is given on each missing concept relation. Finally, all learners can access instructional content of the PDRLS, yet only registered learners can interact with instructional materials in CMFL.

In the following section, we will explore how students can interact with the CMFL website on touch screen devices for the purpose of formative assessment.

3. How Concept Maps for Learning Website Works on Mobile Devices

Initially, a panel of experienced statistics educators created a list of 6 concepts related to ANOVA. The "experts" then rated the degree of relatedness of all pairwise combinations of these 6 concepts, after which the ratings were converted into a knowledge map through the website (which uses the PFNET $(\infty, n-1)$ algorithm to transform relatedness ratings into knowledge maps). The panel of experts then created written explanations, and assembled various instructional materials, to illustrate how any linked concepts in the expert knowledge map are related. These explanations and materials were then uploaded into the website and linked to the expert knowledge map. The resultant expert knowledge map is shown in Figure 6.
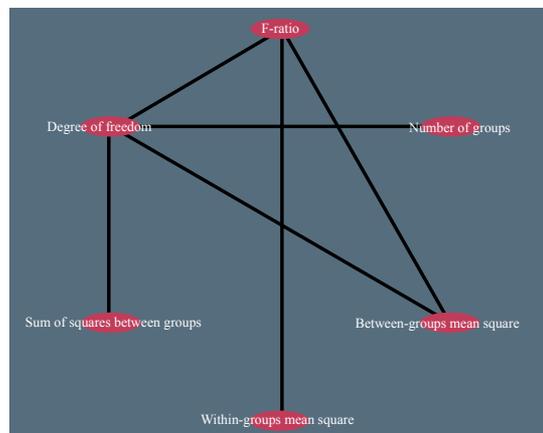

FIGURE 6: An example of expert knowledge map.

At this point, students are ready to use the website to assess their conceptual understanding of ANOVA and receive feedback and remedial instruction. First, students are asked to rate the relatedness degree of concept pairs. Next, they receive visual feedback in the form of their knowledge map superimposed over the expert knowledge map (see Figure 7). A black line appears provided that there is a link between two concepts in both an expert's map and the student map (ex. Between Group Means Square and Fixed Factors). This type of link will be referred to as a relevant link. A grey dotted line appears if there is a link between two concepts in the student map, but not in the expert's map (ex. Between Group Means Square and Deviation of Means from Group Means). This is referred to as an extraneous link. Finally, a red dashed line appears (ex. Between Group Means Square and Random Factors) if there is a link between two concepts in the expert knowledge map, but there is no link between these concepts in the student map. Such links are referred to as missing links.
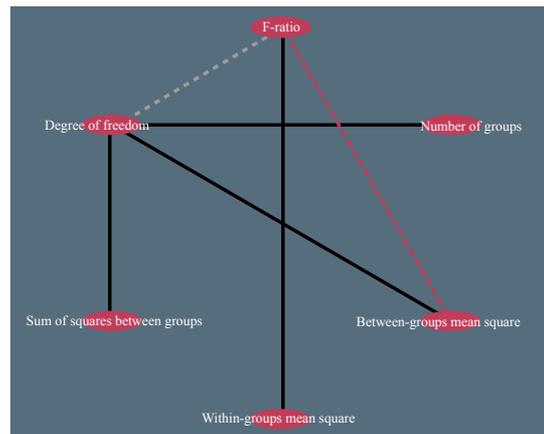
FIGURE 7: An example of visual feedback map.

In addition to this visual feedback, students are given additional instructional materials which are linked to any missing links. First, when students tap on a missing link, a text message appears which explains how the associated concepts are related (see Figure 8); these explanations have been provided by subject matter experts, but can be modified by individual teachers using the website.

Second, if students hold down the mouse/their finger on a missing link, they are able to access linked instructional material intended to illustrate the ways in which the associated concepts are related (e.g., videos, website links, examples, problems, etc.). Again, this material has been provided by subject matter experts, but additional material can be added by individual teachers. As seen in Figure 9, the short explanation is also provided regarding how students should interact with given linked feedback.
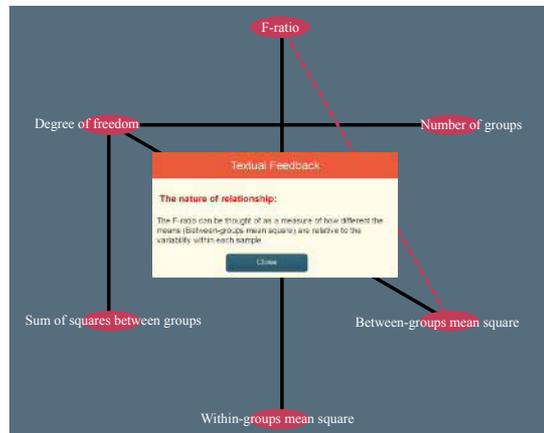
FIGURE 8: An example of given textual feedback.
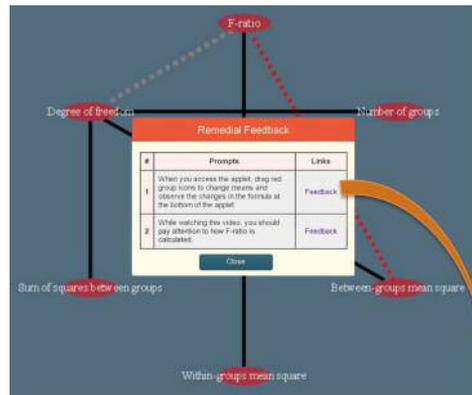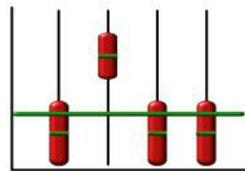
FIGURE 9: An example of given remedial feedback.

4. Conclusion

The aim of this paper is to demonstrate how the Pathfinder Scaling Algorithm converts pairwise ratings into a knowledge map and to present how this algorithm has been used to develop the CMFL website based on the principles of effective formative assessment. In this regard, the Pathfinder structural assessment technique was explained. The fundamental principle of this technique is to compare students' knowledge maps with an expert knowledge map. As a result of this comparison, Pathfinder software computes three different measures showing how students' knowledge map and the referent expert knowledge map compare to each other. However, specific remedial feedback is not given to students by the Pathfinder software. Consequently, software developers have been developing a few digital knowledge mapping tools (e.g., PDRLS and the CMFL website) which use the Pathfinder Scaling Algorithm as a knowledge election technique, but add in an individualized feedback and remediation component.

In the future, we are planning to create more expert knowledge maps on a wider range of subjects (e.g., history, physics and chemistry) to examine the usefulness and effectiveness of the CMFL website across these subjects.

References

Acton, W. H., Johnson, P. J. & Goldsmith, T. E. (1994), 'Structural knowledge assessment: Comparison of referent structures', *Journal of Educational Psychology* **86**(2), 303–311.

Ausubel, D. P. (1978), 2 edn, Holt McDougal, New York.

Boldt, M. (2001), 'Assessing students' accounting knowledge: A structural approach', *Journal of Education for Business* **76**(5), 262–269. doi: 10.1080/08832320109599646.

Boring, R. L. (2005), The validity of human and computerized writing assessment, *in* 'Proceedings of the Human Factors and Ergonomics Society Annual Meeting', Vol. 49, Sage Publications, pp. 759–763.

Casas-García, L. M. & Luengo-González, R. (2013), 'The study of the pupil's cognitive structure: The concept of angle', *European Journal of Psychology of Education* **28**(2), 373–398.

Chen, C. (1998), 'Generalised similarity analysis and pathfinder network scaling', *Interacting with Computers* **10**(2), 107–128.

Chen, C. (2004), 'Searching for intellectual turning points: Progressive knowledge domain visualization', **101**(14), 5303–5310.

Chen, L. H. (2011), 'Enhancement of student learning performance using personalized diagnosis and remedial learning system', *Computers & Education* **56**(1), 289–299. doi: 10.1016/j.compedu.2010.07.015.

Clariana, R. B. (2010), Deriving individual and group knowledge structure from network diagrams and from essays, *in* 'Computer-Based Diagnostics and Systematic Analysis of Knowledge', Springer, pp. 117–130.

Clariana, R. B. & Koul, R. (2004), A computer-based approach for translating text into concept map-like representations, *in* 'Proceedings of the First International Conference on Concept Mapping', pp. 14–17.

Clariana, R. B. & Wallace, P. (2007), 'A computer-based approach for deriving and measuring individual and team knowledge structure from essay questions', *Journal of Educational Computing Research* **37**(3), 211–227.

Cooke, N. J. (1992*a*), 'Eliciting semantic relations for empirically derived networks', *International Journal of Man-Machine Studies* **37**(6), 721–750.

Cooke, N. J. (1992*b*), 'Predicting judgment time from measures of psychological proximity', *Journal of Experimental Psychology: Learning, Memory, and Cognition* **18**(3), 640–653.

Cooke, N. J. (1999), Knowledge elicitation, *in* 'Handbook of Applied Cognition', Wiley, New York, pp. 479–510.

Cooke, N. J., Neville, K. J. & Rowe, A. L. (1996), 'Procedural network representations of sequential data', *Human Computer Interaction* **11**(1), 29–68.

Cooke, N. J. & Schvaneveldt, R. W. (1988), 'Effects of computer programming experience on network representations of abstract programming concepts', *International Journal of Man-Machine Studies* **29**(4), 407–427.

Cooke, N. M., Durso, F. T. & Schvaneveldt, R. W. (1986), 'Recall and measures of memory organization', *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**(4), 538–549. doi: 10.1037/0278-7393.12.4.538.

Davis, M. A., Curtis, M. B. & Tschetter, J. D. (2003), 'Evaluating cognitive training outcomes: Validity and utility of structural knowledge assessment', *Journal of Business and Psychology* **18**(2), 191–206.

Day, E. A., Arthur Jr, W. & Gettman, D. (2001), 'Knowledge structures and the acquisition of a complex skill', *Journal of Applied Psychology* **86**(5), 1022–1033.

DeChurch, L. A. & Mesmer-Magnus, J. R. (2010), 'Measuring shared team mental models: A meta-analysis', *Group Dynamics: Theory, Research, and Practice* **14**(1), 1–14.

Dicerbo, K. E. (2007), 'Knowledge structures of entering computer networking students and their instructors', *Journal of Information Technology Education* **6**, 263–277.

Filiz, M., Trumpower, D. & Atas, S. (2012), Analysis of how well a concept mapping website conforms to principles of effective assessment for learning, *in* 'Concept Maps: Theory, Methodology, Technology. Proceedings of the 5th International Conference on Concept Mapping', Malta, pp. 169–172.

Filiz, M., Trumpower, D. & Atas, S. (2013*a*), The contributions of digital concept maps to assessment for learning practices, *in* 'Cognition and Exploratory Learning in Digital Age', Texas, USA.

Filiz, M., Trumpower, D. & Atas, S. (2013*b*), How a computer-based concept mapping can be used in statistics education: A case of ANOVA, *in* 'United States Conference on Teaching Statistics 2013', North Carolina, USA.

Goldsmith, T. E., Johnson, P. J. & Acton, W. H. (1991), 'Assessing structural knowledge', *Journal of Educational Psychology* **83**(1), 88–96.

Gomez, R. L., Hadfield, O. D. & Housner, L. D. (1996), 'Conceptual maps and simulated teaching episodes as indicators of competence in teaching elementary mathematics', *Journal of Educational Psychology* **88**(3), 572–585. doi: 10.1037/0022-0663.88.3.572.

Guerrero-Bote, V., Zapico-Alonso, F., Espinosa-Calvo, M., Gomez-Crisostomo, R. & Moya-Anegon, F. (2006), 'Binary pathfinder: An improvement to the pathfinder algorithm', *Information Processing and Management* **42**(6), 1484–1490. doi: 10.1016/j.ipm.2006.03.015.

Ichino, M. & Yaguchi, H. (1994), 'Generalized Minkowski metrics for mixed feature-type data analysis', *Systems, Man and Cybernetics, IEEE Transactions on* **24**(4), 698–708.

Kim, M. (2013), 'Concept map engineering: Methods and tools based on the semantic relation approach', *Educational Technology Research and Development* **61**(6), 951–978.

Kim, M. K. (2012), 'Cross-validation study of methods and technologies to assess mental models in a complex problem solving situation', *Computers in Human Behavior* **28**(2), 703–717.

Kivlighan, D. M. & Tibbits, B. M. (2012), 'Silence is mean and other misconceptions of group counseling trainees: Identifying errors of commission and omission in trainees' knowledge structures', *Group Dymanmics: Theory, Research and Practice* **16**(1), 14–34.

Koul, R., Clariana, R. B. & Salehi, R. (2005), 'Comparing several human and computer-based methods for scoring concept maps and essays', *Journal of Educational Computing Research* **32**(3), 227–239.

Lau, W. W. & Yuen, A. H. (2010), 'Promoting conceptual change of learning sorting algorithm through the diagnosis of mental models: The effects of gender and learning styles', *Computers & Education* **54**(1), 275–288.

McGaghie, W. C., McCrimmon, D. R., Mitchell, G., Thompson, J. A. & Ravitch, M. M. (2000), 'Quantitative concept mapping in pulmonary physiology: Comparison of student and faculty knowledge structures', *American Journal of Physiology - Advances in Physiology Education* **23**(1), 72–81.

Moni, R. W., Beswick, E. & Moni, K. B. (2005), 'Using student feedback to construct an assessment rubric for a concept map in physiology', *American Journal of Physiology - Advances in Physiology Education* **29**(4), 197–203. doi: 10.1152/advan.00066.2004.

Nash, J. G. & Nash, J. M. (2003), 'A structural representation of migraine diagnostic criteria: The experts' view', *Headache: The Journal of Head and Face Pain* **43**(4), 322–329.

Rowe, A. L., Cooke, N. J., Hall, E. P. & Halgren, T. L. (1996), 'Toward an on-line knowledge assessment methodology: Building on the relationship between knowing and doing', *Journal of Experimental Psychology: Applied* **2**(1), 31–47. doi: 10.1037/1076-898X.2.1.31.

Sarwar, G. S. (2012), Comparing the Effect of Reflections, Written Exercises, and Multimedia Instruction to Address Learners' Misconceptions Using Structural Assessment of Knowledge, PhD thesis, University of Ottawa.

Schvaneveldt, R. W., ed. (1990), *Pathfinder Associative Networks: Studies in Knowledge Organization*, Ablex Pub, Norwood, N.J.

Taricani, E. & Clariana, R. (2006), 'A technique for automatically scoring open-ended concept maps', *Educational Technology Research and Development* **54**(1), 65–82.

Trumpower, D. L. & Sarwar, G. S. (2010), 'Effectiveness of structural feedback provided by pathfinder networks', *Journal of Educational Computing Research* **43**(1), 7–24.

Trumpower, D. L., Filiz, M. & Sarwar, G. S. (2014), Assessment for learning using digital knowledge maps, *in* 'Digital Knowledge Maps in Education', Springer, pp. 221–237.

Von Minden, A. M., Walls, R. T. & Nardi, A. H. (1998), 'Charting the links between mathematics content and pedagogy concepts: Cartographies of cognition', *Journal of Experimental Education* **66**(4), 339–358.

Wilson, J. M. (1998), 'Differences in knowledge networks about acids and bases of year-12, undergraduate and postgraduate chemistry students', *Research in Science Education* **28**(4), 429–446.