Concept Maps: Theory, Methodology, Technology
Proc. of the First Int. Conference on Concept Mapping
A. J. Cañas, J. D. Novak, F. M. González, Eds.
Pamplona, Spain 2004

'BUT IS OUR CONCEPT MAP ANY GOOD?': CLASSROOM EXPERIENCES WITH THE REASONABLE FALLIBLE ANALYSER

Tom Conlon. School of Education, University of Edinburgh, UK
Email: tom.conlon@ed.ac.uk

Abstract: Classroom concept mapping presents teachers with the problem of how to arrange that learners get feedback on the quality of their concept maps. To address this problem, a software analyser (named the RFA) has been developed that generates scores and hints for student maps by comparing them to an expert map. The analyser is fallible: many of its assessments will be initially incorrect. But it is also reasonable: the student is able to argue for a more favourable assessment. This paper reports classroom trials which indicate that high school students' experience of concept mapping is enhanced by the RFA. Students enjoy arguing with the system, accept its scoring as fair, welcome its hints, and are frequently stimulated to revise their maps to accommodate the feedback obtained.

1 Introduction

Envisage a high school classroom in which students are nearing the end of their work on some topic. In pairs, they have been busily making concept maps to summarise their understanding of the material. One pair has nearly finished as the teacher approaches. She says something encouraging but the students are not quite satisfied with this level of reaction. They want to know: 'But is our concept map any good?'.

The students' question is a reasonable one but it presents the teacher with a problem. Her response could be any one of at least three kinds:

- She could sit down with the students and study and discuss their map. Potentially she can provide rich feedback based on its contents, bringing to bear an own expert understanding of the topic as well as detailed knowledge of her students' learning.
- She could suggest that students self-assess their own map, perhaps by discussion with classmates in which maps are compared with a view to agreeing on improvements.
- She could say 'Your concept map doesn't matter, it's what you learned in making it that counts. Besides, everybody's understanding is unique. So how can a concept map be objectively assessed?'.

The first response is ideal but also perhaps, idealistic. With this approach the teacher gains knowledge of her students' progress and the students benefit from high quality feedback. Unfortunately, since classes normally contain many pupils it is feasible only seldom. The second response might be productive in some contexts but it is tricky since it implies a possibly lengthy process of interaction between peers whose ability to provide useful assessment is uncertain. As for the third response, it seems unhelpful and misguided. Of course the process of concept mapping is more important than the product and it is true that a map provides a personal view of a domain. But surely, a summary of routine curriculum material ought to capture key features in more or less recognisable ways. Feedback plays a vital role in learning generally and to disregard the possibilities that concept maps afford for feedback is to miss an opportunity and also, possibly, to undermine students' confidence in concept mapping.

The development of the Reasonable Fallible Analyser (RFA) was motivated by a concern that in practice, many students are making concept maps for which they get little or no feedback. The RFA is a computer program that compares a student's concept map to one produced by an expert or able peer and produces a score along with hints for improvement. Because it imposes no restrictions on students' language, the analyser is fallible: many of its assessments will be initially incorrect. But it is also reasonable: the student is able to argue for a more favourable assessment.

The remainder of the paper is in three main parts. First, a brief system description of the RFA is offered. Second, some classroom experiences are reported that are based on trials with the RFA in an Edinburgh high school. The trials indicate that high school students' experience of concept mapping is enhanced by the RFA. Students enjoy arguing with the system, accept its scoring as fair, welcome its hints, and are frequently stimulated to revise their maps to accommodate the feedback obtained. Third, some discussion is offered around the content of this work, including its relationship to the wider context of assessment.

2 The Reasonable Fallible Analyser

The user interface to the RFA is shown in Figure 1 below. The design of the system has been reported elsewhere (Conlon 2004) and only a brief summary is given here. In contrast to other concept map analysers (Biswas et al 2001, Chang et al 2001), the RFA aims to be extremely flexible and unobtrusive upon the normal concept map building process. Thus a first design goal was for avoidance of restriction to particular subject domains, vocabularies, or even map building environments — the RFA (when armed with a suitable comparison map) should be able to analyse any map on any topic, implemented in any software environment, using language freely chosen by the student. This implies fallibility on the system's part, because unrestricted concept maps will sometimes contain text that the analyser cannot reliably interpret. So a second design goal was for openness: the system should be able to negotiate its assessment with the student. Openness is not simply a tactic to compensate for the system's lack of ability, however. It also reflects a promotion of the learner to a status of sharing responsibility for the learning (and assessment) process. This is consistent with recent thinking (discussed later) about formative assessment in education and also with trends in the design of computer-based learning environments (Morales et al 1999). It seems particularly appropriate for concept mapping which from the start has been motivated by the principle of 'learning how to learn' (Novak and Gowin, 1984).

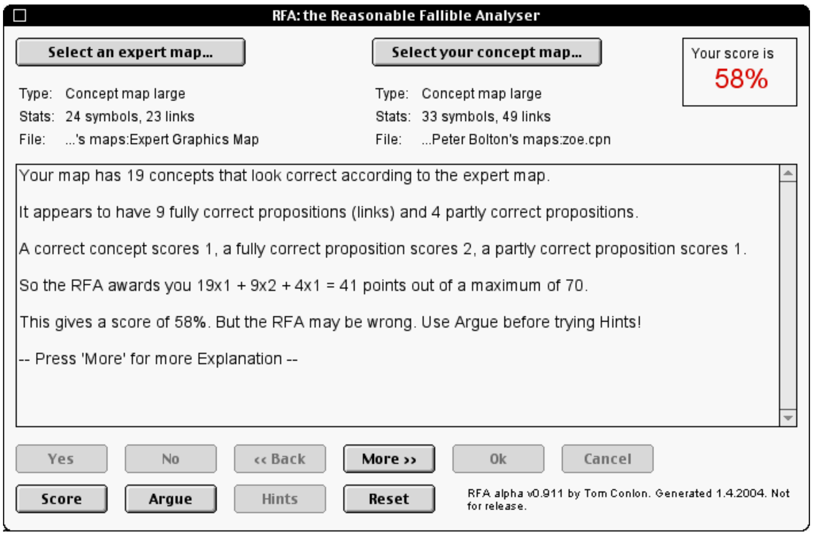

Figure 1 User interface to the RFA

In a little more detail, the analyser was designed to satisfy the following criteria:

- The system should be interoperable with a variety of concept map building tools. At present however it accepts only maps in the format of Conception (www.parlog.com).
- No special pre-processing ought to be necessary to enable any concept map to serve as an expert map.
- The system ought not to depend upon assumptions about the content of student maps. In particular, maps are not expected to be constrained to a pre-specified lexicon.
- Scoring of student maps should use a formula that credits both correct concepts and correct propositions, as specified by the comparison map. Figure 1 indicates the formula that is currently implemented; of course, other formulae are possible. A transparent explanation for scores should be provided.
- When the system detects that a concept or proposition mismatches anything in the comparison map, the student ought to have the opportunity (via an 'Argue' function) to claim that an equivalence was intended. Thus, the initially computed score should be regarded as provisional and capable of being upgraded by negotiation.
- The system should be able to offer hints relating to missing concepts, missing links, and dubious relationships. Recommendations to add new concepts should be limited to those that can be attached to concepts that are actually present in the student's map.

The RFA is implemented in LPA Prolog++ (www.lpa.co.uk) and runs under both Windows and Macintosh operating systems. To boost its general language capabilities the system incorporates WordNet, a lexical database accessible from Prolog that incorporates a dictionary of 200,000 words (Witzig 2003). In addition, a separate User Dictionary is incorporated which can be edited to define specialist synonyms that would not be expected to occur in WordNet.

As Figure 1 indicates, when a map score has been initially computed the student is invited to negotiate an improved score by means of an 'Argue' button. Three types of argument are then possible. One is a protracted system-led interaction in which the student is asked to confirm or deny the presence of possible defects that have been detected in the map. This option is useful for novices but perhaps frustrating for experienced users who should be able to infer (from the score explanation, which can be viewed in far more detail than the screen shot shows) in what way the RFA's analysis is erroneous. Therefore two other types of argument are available which enable the student to take more initiative in the dialogue, by identifying respectively specific concepts or propositions that have been misconstrued by the system.

3 Classroom Trials

When a limited prototype version of the RFA was tested under laboratory conditions with a sample of student concept maps, performance seemed promising. The RFA's scores correlated quite closely with scores calculated manually by an expert human assessor. The system's pre-argument scores under-rated student maps by an average of 25% but the argument process reduced the difference to an average of 16%. Thus the RFA demonstrated its fallibility, but also its reasonableness in adjusting its assessments in the direction of greater accuracy in response to argument.

Classroom trials with the RFA were undertaken in an Edinburgh high school in March 2004. The trials were intended to answer some basic questions. Did students find the analyser usable? Was it easy to learn, effective in providing fair and helpful feedback, and enjoyable to use? Of particular concern was students' perceptions of the analyser's fallibility and its provision of an argument function. Would they take this function seriously or perhaps, try to 'cheat' by claiming as synonyms terms that were not genuinely synonymous? In addition, although it was not attempted at this stage to quantify the effect on learning, the effect of the RFA on the overall concept mapping process was of interest. Would students be stimulated by interaction with the RFA to revise their concept maps?

During the trials, the RFA was made available to 40 students aged 14-16 over a period of two weeks. The students were following courses in computer technology at a variety of academic levels. The class teacher, who regularly incorporated concept mapping activities within his teaching, had students engaged in the preparation of concept map summaries of course material using Conception (www.parlog.com) software. This much was more or less conventional practice. The novel feature was that students were told that on these occasions, it would be possible to use the RFA to get feedback on their maps. Instruction on the RFA was minimal: students were issued with a one-page handout that summarised the operation of the user interface. Expert maps had been prepared in advance by the class teacher and their file names were displayed on the class whiteboard. No use was made of the RFA's capability to accommodate specialist vocabulary within the User Dictionary.

The classroom trials produced data from three main sources: direct observation by the class teacher and the researcher; log files generated automatically by the RFA that (unbeknown to students) recorded details of user interaction; and a questionnaire that was completed by students immediately after the lessons. The questionnaire elicited responses to a set of provided statements using a four-point Likert scale: 1 = Disagree strongly, 2 = Disagree, 3 = Agree, 4 = Agree strongly. For summary purposes, it will be convenient below to structure the results around categories of questions from the questionnaire, supplementing these with evidence from direct observation and log files where appropriate.

3.1 Attitudes to Concept Mapping

Clearly, students' attitudes to the RFA need to be interpreted in the context of their attitudes to concept mapping generally. These attitudes are summarised by the questionnaire responses shown in Table 1 where the percentages represent the total of students who answered either 'Agree' or 'Agree strongly' with the statement. As can be seen, students are largely positive towards concept mapping and they favour computer tools for map construction. An analysis by gender revealed that boys more frequently than girls identified concept mapping as hard work — 37% of boys compared to 17% of girls agreed with this statement. However, girls were heavily outnumbered in our sample (only 12 females out of 40 students) and neither this nor any other gender difference reported here is statistically significant.

Statement	Agreement
Concept mapping helps me to learn.	95%
Concept mapping is hard work.	31%
If I had the choice, I wouldn't make concept maps again in future.	23%
I prefer paper-and-pencil to a computer for concept mapping.	15%

Table 1. Students' attitudes to concept mapping

3.2 Attitudes to Feedback

As mentioned previously, a hypothesis that motivated the development of the RFA was that students generally are making concept maps for which they get little or no feedback. This hypothesis was tested by the two questionnaire statements shown in Table 2. The lack of feedback is confirmed. Furthermore, it matters: a very large majority of students does care about knowing whether their maps are good or bad.

Statement	Agreement
Until now, when I have constructed a concept map I haven't known whether it's good or bad.	69%
It's important to me to know whether my concept map is good or bad.	92%

Table 2. Students' attitudes to feedback

3.3 RFA as a Stimulus to Map Revision

The class teacher encouraged students to use the results of each interaction with the RFA as a stimulus to revise their maps. In practice, students were observed to do this often and it was clear that they were motivated in their map-making by the prospect of submitting their maps to the RFA for analysis. As Table 3 shows, students agreed that the RFA was effective in this role.

Statement	Agreement
Using the Analyser made me want to redo parts on my concept map.	87%
I would put more effort into my concept maps in future if I knew they were going to be scored by the Analyser.	90%
Using the Analyser made me realise that I need to learn more about the subject matter of my map.	79%

Table 3. The RFA as stimulus to map revision

3.4 Attitudes to the Argument Function

Observation revealed that students found the RFA easy to use. The '10-minute rule' proposed by Nelson (1980), whereby the ability of a novice to be able to learn to use a new computer system within 10 minutes is taken as the criterion of whether the system is easy to learn, seems to be broadly met by the RFA. However, it was evident that students sometimes skipped over the text that provided detailed explanation of how their maps had been scored and they often showed hesitancy when selecting between different options of the Argument function. Nevertheless, students were quite tenacious in argument and they commonly achieved gains of around 20% over the system's pre-argument scores. Most of these gains were due to the student's identification of mismatches between the textual rendering of concepts in the two maps: in every case the mismatch was plausible, that is, no cheating was encountered.

Students' questionnaire responses (Table 4) confirm the evidence of classroom observation that they generally enjoyed arguing with the system, accepted its scoring as fair, and were basically honest in their approach. In fact, students were surprisingly satisfied even by the RFA's pre-argument scores even although these typically undervalued their maps considerably. The belief on the part of 44% of students that cheating would necessarily be detected by the system is misguided, of course. In fact, the RFA is very gullible. This is not a point that was stressed in class and it remains to be seen how students' attitudes to the system may change as increased exposure (presumably) reveals its gullibility.

A couple of gender differences in this area are suggestive. First, girls were more satisfied with the scores their maps were awarded than boys: 92% of girls regarded the pre-argument scores as fair and for post-argument scores this rose to 100%. The corresponding figures for boys were 63% and 96%. Second, girls were more likely to believe that cheating would be detected: 58% of girls expressed this view as compared to 37% of boys.

Statement	Agreement
It was fun to use the 'Argue' button of the Analyser.	85%
The score that was given to my map(s) by the Analyser even before I used the 'Argue' button was fair.	72%
The score that was given to my map(s) after I used the 'Argue' button was fair.	97%
It was frustrating to argue with the Analyser.	23%
I was keen to try to raise my map's score by arguing with the Analyser.	92%
When arguing with the Analyser, I was tempted to cheat by claiming that different words meant the same thing.	28%
If I tried to cheat when arguing with the Analyser, it would be certain to find me out.	44%
When arguing with the Analyser, I always did my best to answer its questions as honestly as possible.	90%

Table 4. Attitudes to the Argument function

3.5 Perceptions of Scores and Hints

As Table 5 shows, students generally expressed approval of the RFA's score explanations and hints. For the former, a higher proportion of girls (92%) was satisfied than boys (63%). More than three-quarters of students claimed that they would like to use the RFA with all their future map-making activities. The class teacher agreed that the RFA's hints were well-founded and likely to be helpful to his students' learning. The system's ability to generate a Certificate summarising these hints, which students could print and take away, was especially welcomed (for an example see Figure 2).

Statement	Agreement
The Analyser is good at explaining how it calculates its map scores.	72%
The hints provided by the Analyser were helpful.	85%
If I had the choice, I would use the Analyser every time I made a map.	77%

Table 5. The RFA as stimulus to map revision

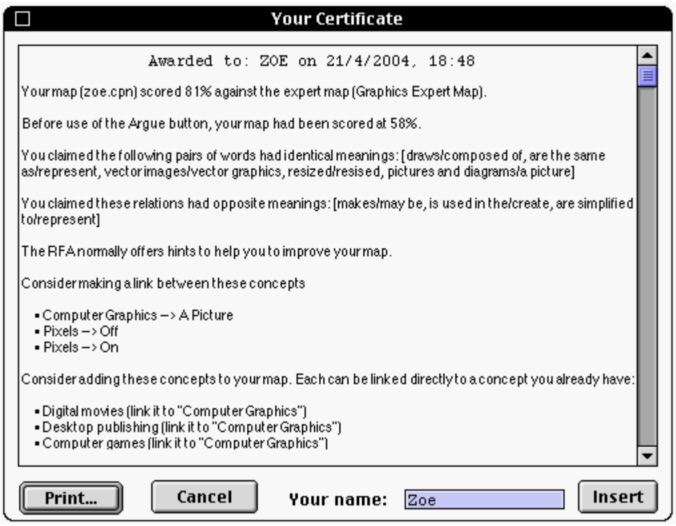

Figure 2 A post-argument certificate generated by the RFA

4 Discussion

In this section two central issues are discussed. First, the RFA is related to the wider context of assessment in education. Second, the scope of use of the RFA and its appropriate associated pedagogy are considered.

4.1 Formative Assessment

A conventional distinction made in the literature is between formative and summative assessment. Briefly, formative assessment aims to produce information which helps students to improve their learning. The information may be directed at the learner (formative self-assessment) or at the teacher (diagnostic assessment) or possibly, at both. Summative assessment aims to produce information which summarises what the student has learned, for example for the purpose of reporting to parents or awarding academic credits. Primarily, the RFA is intended to contribute *formative self-assessment*. It has potential in diagnostic assessment, even in its present form, and it should be straightforward to develop (for instance) a variant of the system that can batch-process a class set of concept maps and generate a profile of represented concepts and relationships which can guide future teaching action.

The development of a new technology for formative assessment of concept maps is timely. In Scotland and perhaps in other countries too, education policy makers now urge schools to make more use of formative assessment (Simpson 2003). The reasons are several and include awareness of the limitations of summative assessment, the need to promote self-monitoring, self-motivation and self-reliance among learners, and a general desire to shift from a teacher-centred, transmissional pedagogy towards a learner-centred, constructivist one. The effectiveness of concept mapping for learning has been demonstrated many times but its potential contribution to formative assessment seems to have been neglected. Summative assessment, on the other hand, has received a lot of attention from concept map researchers (e.g. Rice et al 1998, Ruiz-Primo et al 1998, West et al 2002). This seems odd, not only because concept mapping is most naturally seen as a developmental tool but also because formative assessment generally is much less beset by the reliability and validity issues which have predictably dominated discussion about the role of concept mapping in summative assessment.

4.2 Scope and Pedagogy

The scope of the RFA (by which is meant, the range of contexts in which its use is appropriate) and its associated pedagogy need to be clarified by further research. However, since the system depends upon the availability of a comparison map it seems likely that exploratory forms of concept mapping (for example, mapping out a plan for a piece of creative writing) are outwith the RFA's scope. However, much classroom concept mapping seems to be about *summarising* rather than exploring. Commonly, students' maps are abstractions of routine curriculum subject matter. For such tasks, use of the RFA should be feasible.

A pedagogy that incorporates the RFA, and which generalises the approach used in the classroom trials, is illustrated in Figure 3. Here, labels represent processes to be accomplished by the student. Although more investigation is needed, it seems predictable that the processes 'Review domain' and 'Discuss with teachers/peers' are likely to be crucial to ensuring that the RFA's feedback is productive within a cycle of meaningful learning. Without them, there is a danger that some students will engage in a shallow edit-submit-edit cycle that utilises feedback only at surface level and which limits the scope of learning to the names of concepts and relations. Such students might succeed eventually in bringing their map into alignment with the expert map but without much change in their own personal conceptual frameworks.

The scaffolding (support) which the teacher supplies initially with the task will also be important. A study by (Chang et al.2001) compared experimentally two groups of high school students who were concept mapping in the domain of biology. One group was scaffolded initially by the provision of a partly completed map whilst the other group built maps from scratch. Both groups used the same concept mapping environment that provided feedback based on an expert map and which imposed (even for the 'scratch' students) a prespecified list of concept and relation names. Tests of learning showed significantly better results for the scaffolded group.

Generally in Scotland, skilled classroom teachers acknowledge that the extent and form of scaffolding is a key factor in learning and that it is often necessary to adjust scaffolding so as to differentiate between learners at different stages of development. In classroom trials with the RFA, scaffolding was provided by the class teacher by eliciting a partial list of concept names through whole-class brainstorming prior to the commencement of concept mapping. If Chang's study is a guide, it would be useful to consider the option of making available (to some students) a partially completed map file. On the other hand, the incorporation into pedagogy of the RFA might enable students to succeed with tasks that are more open-ended: their initial maps may be wide of the mark but with feedback they can be improved.

It remains to be seen how attractive will be this pedagogy to teachers. There is some evidence that the uptake of concept mapping by high school teachers is low (Conlon & Bird 2004). No doubt several reasons can be found for this but among them is probably the fact that concept mapping presents teachers with the problem of how to arrange that learners get feedback on the quality of their maps. The RFA addresses this problem and adds to concept mapping some additional benefits that even traditionally minded teachers should appreciate: a clear goal for students who are motivated to submit their maps to the system; an enjoyable and productive use of new technology; and a liberation of the teacher's time that makes room for individualised coaching and other creative teaching activities.

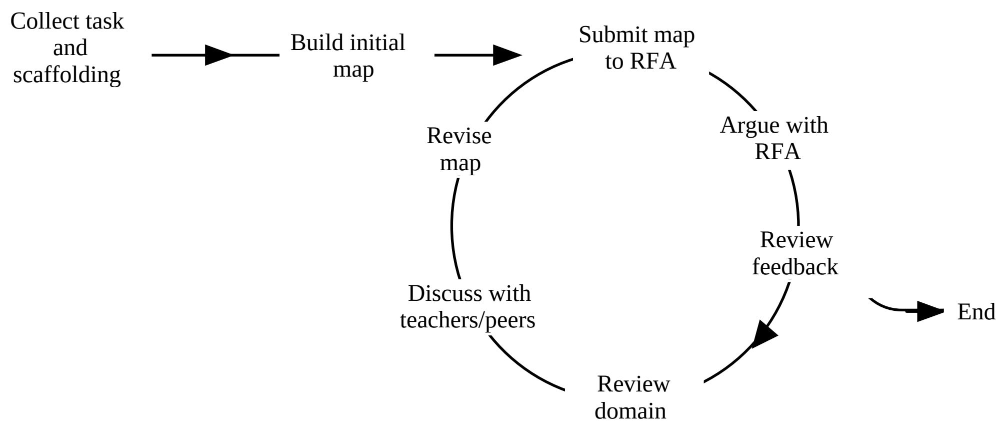

Figure 3 A pedagogy for concept mapping with the RFA

5 Conclusion

There is clear scope for development of the RFA. In particular, provision could be made to enable the system to operate more incrementally, that is, as coach rather than critic; the argument function could be enriched; and the system could be extended to exploit a multiplicity of comparison maps on a given domain, rather than relying on only a single comparison map.

Such developments however will need to be guided carefully by theories and empirical trials. The trials reported in this paper indicate that even in its present form, high school students' experience of concept mapping is enhanced by the RFA. Students enjoy arguing with the system, accept its scoring as fair, welcome its hints, and are stimulated to revise their maps to accommodate the feedback obtained. Because the RFA enables formative assessment of students' concept maps without imposing any restrictions upon their freedom of expression, and because of its flexibility in adapting to each new domain merely by the addition of a comparison map, it is claimed that this technology and its associated pedagogy represent a significant contribution to concept mapping.

6 Acknowledgements

Thanks are due to Stephen Docherty and Adam Caldwell for their help with the classroom trials. I am very grateful also to the anonymous CMC2004 referees for their valuable comment and encouragement.

7 References

Biswas, G., Schwatz, D. & Bransford, J. (2001). Technology Support for Complex Problem Solving. In: Forbus, K. & Feltovich, P. *Smart Machines in Education.* MIT Press.

Chang, K., Sung, Y & Chen, S. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning* 17, 21-33.

Conlon, T. & Bird, D. (2004). Not Yet Within the Mainstream: Concept Mapping in a Scottish High School. *Proceedings of CMC2004.* Pamplona, Spain, September 14-17.

Conlon, T. (2004). 'Please Argue, I Could Be Wrong': a Reasonable Fallible Analyser for Student Concept Maps. *Proceedings of Ed-Media 2004,* World Conference on Educational Multimedia, Hypermedia, and Telecommunications, Lugano, Switzerland, June 21-26, 2004.

Morales, R. (1999). *Proceedings of the Workshop on Open, Interactive and other overt approaches to learner modelling.* AIED99: World Conference on Artificial Intelligence in Education, Le Mans, France.

Nelson, T. (1980). Interactive Systems and the Design of Virtuality. *Creative Computing*, Nov-Dec.

Novak, J. and Gowin, B. (1984). *Learning How to Learn.* Cambridge University Press.

Rice, D., Ryan, J., & Samson, S. (1998). Using Concept Maps to Assess Student Learning in the Science Classroom: Must Different Methods Compete*? Journal of Research in Science Teaching* Vol 35 No 10 pp1103-1127.

Ruiz-Primo, M., Schultz, S., Li, M. & Shavelson, R. (1998). *Comparison of the Reliability and Validity of Scores From Two Concept Mapping Techniques.* CSE Technical Report 492, Graduate School of Education & Information Studies, University of California, Los Angeles.

Simpson, M. (2003). Diagnostic and Formative Assessment in the Scottish Classroom. In Bryce, T. & Humes, W. (Eds) *Scottish Education.* Second Edition: Edinburgh University Press. pp721-730.

West, D., Park, J., Pomeroy, J., & Sandoval, J. (2002). Concept mapping assessment in medical education: a comparison of two scoring systems. *Medical Education* 36: 820-826.

Witzig, S. (2003). *Accessing WordNet from Prolog.* Artificial Intelligence Centre, University of Georgia.