CHANGES IN STUDENT KNOWLEDGE STRUCTURES IN SCIENCE:
HOW TO MEASURE STUDENT KNOWLEDGE STRUCTURES

HELENA DEDIC, STEVEN ROSENFIELD AND MIRIAM COOPER, **VANIER COLLEGE AND CONCORDIA UNIVERSITY'S CENTRE FOR THE STUDY OF CLASSROOM PROCESSES**

SUMMARY

A novel approach to the development of instruments for testing conceptual structures. As we move from fact-based education towards conceptual development, teachers, for whom discipline-specific tests of student recollection of facts or algorithms are routine, must learn to test for conceptual structures as well. This presentation outlines a methodology that makes the development and use of such instruments relatively quick and easy. Thus, their usage could become common practice, encouraging students to focus on understanding, and enabling teachers to estimate the "proximal learning zone".

WHAT IS A KNOWLEDGE STRUCTURE?

If we are to discuss changes in student knowledge structures we must begin with a definition of the term "knowledge structure". We use this term to describe an individual's overall organization of knowledge. This includes both factual information and the links between different pieces of information. A simplistic analogy might be to think of a human knowledge structure as the human equivalent of a computer database, filled with records and structured by multiple indexes that tie records together in different ways and so allow access to records in different ways. We use the term "mental model" to describe a mental representation of a specific object or a concept, *e.g.*, the mental model of a vector, and its connections to other concepts, within a knowledge structure. Simplistically, a mental model is the analogue of a record or entry in the database, along with its links to other records. We refer to changes in the student knowledge structure as conceptual change and so this too requires a definition.

WHAT IS CONCEPTUAL CHANGE?

Rote Learning = No Conceptual Change

When a student encounters new concepts or a new conceptualization of ideas seen before, conceptual change may not take place. The student may simply file the new concept away and make little or no attempt to integrate it with current knowledge. In terms of our database analogy, in this case a new entry or record is created in the database, but few if any indexes are updated, and no new type of indexing is contemplated. Thus, the new information is isolated in the knowledge structure/database and so it will be difficult to access or retrieve this new information unless cued by circumstances virtually identical to those present at the time of storage. As teachers we witness this approach to learning in science all too often. Students drill themselves in examples, and if asked that exact question, present a perfect response. However, if the example is changed even a small amount, so that to the teacher the question may appear unchanged, nevertheless, to the student it is a complete new problem and cannot be solved.

Meaningful Learning = Conceptual Change

In contrast, when the student not only adds the new information to their knowledge structure, but integrates it into the existing structure, or even restructures existing knowledge to accommodate the new knowledge, then conceptual change has taken place. In terms of our database analogy, a new entry is made in the database, most or all indexes are updated to include the new entry, and perhaps new indexes are initiated. Thus, the new information is well connected to older information and can be retrieved in a wide variety of circumstances. Unfortunately, as teachers we witness this type of behaviour less frequently. Students must probe their understanding of new material, seeking to compare and contrast to earlier experience, and actively create links or relationships to earlier understandings.

Our PAREA research project, *Changes in Student Knowledge Structures in Science*, is principally aimed at testing strategies that can be used in the classroom by all teachers to elicit conceptual change in students. However, if one is to measure the effectiveness of such strategies, one must begin by determining student knowledge structures, at least in an area being covered in the classroom. While a knowledge structure is a dynamic thing, the best we can hope for is a static measurement, a snapshot or x-ray of that structure at a given moment in time. In theory at least the tests that we all give in our classrooms attempt to do just this. Unfortunately, research by Halloun and Hestenes (1985) has shown that there is little correlation between scores on such classroom achievement tests and more elaborate tests designed specally to determine student understanding of concepts.

SOME STANDARD TECHNIQUES USED TO CAPTURE KNOWLEDGE STRUCTURES

Structured Interviews

A standard technique used to measure student knowledge structures is the structured interview. While this technique is generally accepted as effective, it is expensive of both time and effort, for both teacher/researcher and student. Thus, experiments that use this technique must be small in size and thereby diminishing the power of their results. Worse, such a technique would be impossible for teachers to use on a regular basis in their classrooms, and so has narrow applicability outside the artificial world of experimentation.

FITS (Fill-In-The-Structure)

In 1994 Naveh-Benjamin & Lin suggested a technique that seemed much more time and cost effective. They present each student with a concept map from which some concepts are missing at different levels of a hierarchical structure, and provide a list of concepts, mixed with distracter concepts. The student task is to fill in the missing concepts from the list. To use this technique in the classroom, one must adapt FITS to the domain taught and then develop evaluation criteria. The evaluation criteria centre around a comparison of student maps and those of the researchers/experts. We had little difficulty with the notion of adapting FITS to the domain of Physics, but our efforts stalled when we attempted to develop evaluation criteria. The measures used by Naveh-Benjamin & Yin were: content similarity, or % of concepts "correctly" placed; hierarchical structure, or the % of concepts placed in the appropriate level of the hierarchy. Our problem was that our expert knowledge structures, as presented in concept maps, were not unique. This made both of the measures somewhat arbitrary.

Pathfinder

In 1981 R. W. Schvaneveldt and F. T. Durso proposed an algorithm for generating concept maps automatically from proximity matrices that encode the distances between concepts. Subsequently, they generated the computer program Pathfinder to automate this process. They, and others, have used this program in a process to measure knowledge structures. They begin by giving students a list of concepts and the students are asked to rank proximity or closeness of the concepts on a scale from 1 to 8. These values are entered into a symmetric matrix whose rows and columns represent the concepts. The Pathfinder software then converts the proximity matrices into concept maps. Subsequently the software analyses the differences between student and expert maps to assess similarity of knowledge structures. Two measures are involved: *C* and *r*. *C* is the degree to which the a node in the expert concept map and in the corresponding node in the student concept map are surrounded by similar neighbourhoods of nodes, averaged across all nodes in the maps. The measure *C* varies from 0 for complementary concept maps to 1 for identical concept maps. The measure *r* is the coefficient of correlation for pair-wise distances between concepts. While we were pleased with the automation of this process, and could easily prepare lists of appropriate concepts, adapting the method to any topic, we were concerned, as with FITS, with the possibilities for disagreements between experts. Further, it seemed to us that student judgements of the ìclosenessî of any two concepts on a scale from 1 to 8 would be somewhat arbitrary. Thus, we struggled to find a way to make judgements clear, with "correct", "incorrect" and "more incorrect" answers.

DEVELOPING OUR OWN "MOTION QUESTIONNAIRE"

Motion Concepts

We began with the notion that we could use some phrases that expressed practical or real life expression of concepts taught in the study of motion in Physics. For example, here are six from our original list of such phrases: A - there is no net force acting on the car; B - there is a net force due North acting on the car; C - a ball dropped from the car falls straight down; D - a ball dropped from the car falls along a parabolic path; E - the car is at rest; F - the car is moving at constant speed due North. Instead of asking students to decide how closely these concepts were related, we decided to have them choose which of four relationships the concepts satisfied. That is, with one of the phrases acting as "statement a" and another as "statement b", students would be asked to choose: 1 = "a" implies "b"; 2 = "a" may imply "b"; 3 = "a" is unrelated to "b"; or, 4 = "a" cannot imply "b". With n phrases considered in all possible pairings, this would generate n^2 questions, and an nxn proximity matrix. We have now been through at least six versions of phrases and different wordings for student choices. While we have no statistical data to report, we do have preliminary observations, which in part explain the multiple iterations.

PRELIMINARY RESULTS

We ran our initial versions of the Motion Questionnaire on small numbers of students selected conveniently as they studied at the Science Resource Centre, that members of the research team staff and run at Vanier College. We immediately discovered that both the number of questions and their repetitiveness caused confusion amongst students.

Thus, we limited the number of concepts addressed at one time, and removed many of the more obvious questions to avoid repetition.

As we tested further, we noticed that subtle differences in the types of relationships offered in the four choices, and simple wording changes caused disproportionate confusion amongst students. We realized that we would have to train, or at a minimum, expose students to this new kind of testing in advance. Further, a number of iterations in wording were required to more clearly express the relationship choices and their directionality.

Even with such attempts, it became clear that there are real problems in student understanding of if ... then type logical statements. To us it appears that we will need to train students in elementary logic.

Despite all of the above problems, by means of post-test interviews, we clearly ascertained that some incorrect answers were due to conceptual difficulties.

CONCLUSIONS

Interviews showed that some of the errors captured were misconceptions, so there is hope that this mode of testing, once mastered, can isolate and identify such misconceptions. Thus, our future plans include:

i. showing more humility and respect for social sci entists who on a regular basis generate measures;
ii. training students in logic and these kinds of ques tions (both of which we feel are desirable on their own);
iii. testing the effectiveness of training;
iv. continuing modification of our questionnaire, run ning more pilot tests accompanied by interviews.

BIBLIOGRAPHY

Goldsmith, T.E., Johnson, P.J. & Acton, W.H. (1996), Assessing Structural Knowledge. *Journal of Educational Psychology*, 83, 88-96.

Halloun, I. A., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics*, 53, 1043-1055.

Naveh-Benjamin, M. & Lin Y.-G. (1994). Measuring and Improving Students' Disciplinary Knowledge Structures. In P. R. Pintrich, Brown, D. R. & Weinstein, C. E. (Eds.), *Student Motivation, Cognition, and Learning: Essays in Honour of Wilbert J. McKeachie.* Hillsdale, NJ: Lawrence.
R. W. Schvaneveldt (1990) (Editor), *Associative Networks: Studies in Knowledge Organization.* Ablex Publishing. Norwood, N.J., ISBN 0-89391-624-2