Survey of Learners' Knowledge Structures: Rationales, Methods and Instruments

Peiling Wang[†], **Stephen Bales, Jason Rieger, Yan Zhang**

School of Information Sciences, College of Communication and Information, The University of Tennessee, Knoxville, 445 Communications Building, 1345 Circle Park, Knoxville, TN 37996-0341 Email: peilingw@utk.edu

Abstract

How human beings organize knowledge and to what extend individuals' organizational structures overlap are important research questions in cognitive science (Lakoff, 1987) as well as in information science (Ingwersen, 1992). Effective IR principles can be derived from the understanding of human knowledge organizational structures. This project explores a field study design to observe learners' organization of concepts of a subject domain during a course of learning. The purpose of the study is two-fold: (1) to investigate how learners build and organize concepts learned along the process and (2) to design data collection methods and instruments. The project is carried out in two introductory information science courses as two studies. The ultimate goal is also two-fold: (1) to facilitate learning in introductory courses by adopting metacognitive tools and (2) to advance information retrieval by incorporating effective knowledge structures into IR systems.

Introduction

Human knowledge may be classified into three types: (1) domain knowledge, (2) task knowledge, and (3) phenomenological knowledge. Domain knowledge is ontological, specific, and typically learned in school as subject matter. Domain-specific knowledge consists of concepts and their relationships, similar to a thesaurus. Task-dependent knowledge is epistemological, strategic and procedural, typically obtained by applying domain knowledge to problem solving in situations. For example, procedural knowledge differentiates experts and novices. Phenomenological knowledge is experiential and includes tacit knowledge (automatic and not readily available to consciousness), episodic knowledge (personal), and socio-cultural knowledge (shared among a culture of people), which are uniquely learned from life (Jonassen, 2000).

Individuals must take an active role in representing and organizing what they learn in order to use knowledge effectively and to better assimilate new information. Ausubel, et al. (1978) argue "The most important factor in learning is what the learner already knows." (p. 163) New information must be assimilated to the current knowledge structure to make sense. Hence the important question: how do human beings organize their knowledge? One assumption is that individuals' knowledge spaces consist of concepts and their relationships. A concept is an idea, a tangible object, an abstract entity, a unit of thoughts, and so on. Novak (1998) defines "*concept* as a perceived regularity in events or objects, or records of events or objects, designated by a label." (p. 21) A label for a concept is a linguistic (or other format) expression of the concept. Concepts are thus wrapped in terms (words or phrases) to form a vocabulary consisting of all the terms that a person knows. Terminology, as opposed to vocabulary, includes only the technical terms defined by a knowledge domain. Relationships between concepts are the glue that holds concepts together to form stable yet dynamic structures.

A popular structural representation of knowledge is hierarchical organization, typically seen in classifications or thesauri, which may be depicted as a graph similar to a reversed tree or organizational chart, starting from the more general concept(s) to more specific ones (a top-down configuration)[1]. Other types of representations are structurally more complex, such as a temporal order similar to a flow chart or an interconnected network similar to a web expanding from a center[2]. Many different representations might co-exist within an individual's memory; different aspects of a specific domain knowledge might be represented through different structures (Rumelhart & Norman, 1985; Jonassen & Grabowski, 1993).

Ingwersen (1992) illustrates that cognitive structures vary across individuals and situations. However, it is the interpersonal similarity in mental contents that enables human communication. The challenge is in observing individuals' knowledge structures to understand differences and similarities. (Lakoff, 1987)

With a focus on domain-specific knowledge, this project is carried out to understand how learners build and revise subject knowledge during a course and to observe the similarities and differences in knowledge structures. The immediate goals of this project are to develop robust methods for empirical studies and to

---

[†] Corresponding author

design instruments for measuring cognitive structures. The ultimate goal of this line of research is to develop theory or models from empirical evidence on the process of knowledge development and cognitive structures of domain knowledge.

We set out to investigate the following research questions:

1. What changes occur in learner *vocabulary* during the learning process in a course?
2. What kinds of *conceptual overlapping* exist among learners?
3. Which semantic relationships between concepts do learners *recognize*?
4. Do *semantic relationships* between concepts change overtime? If yes, what are the *changes*?
5. Do *learning styles* play any role in the development of knowledge structures?
6. Will concept mapping be an effective alternative to traditional exams for *evaluating* learners' domain knowledge and misconceptions?

At the time of this writing, this project has completed two phases in which data collection instruments were tested and revised. This paper reports on the study design, especially the data instruments, and the preliminary results of two studies conducted in 2003.

Related Literature

Using a geographic metaphor, researchers and educators have adopted concept maps to represent knowledge structures stored in human memories. Wandersee (1990) suggests that the graphical depiction of individual knowledge structure (through concept maps, concept circles, and vee diagrams) amounts to a "cartography of cognition," providing a visual summarization of knowledge while aiding continuing exploration, integration of new ideas, and reevaluation of existing mental structures.

Building upon Ausubel's (1963) earlier research concerning knowledge assimilation and meaningful learning, Novak and his associates (Ausubel, Novak & Hanesian, 1978; Novak & Gowin, 1984; Novak, 1990; 1998) developed and advocate concept mapping for science education as a metacognitive tool. Concept maps provide a schematic summary of what has been learned (including misconceptions) and make clear to both students and instructors the key ideas that they must focus on for any specific learning task. Researchers in education have come to a general consensus in favor of the use of concept mapping in the classroom to encourage meaningful learning while providing insight into the development of student knowledge structure (and its relationship to "expert" referents). Subsequently, concept-mapping exercises, often articulated by means of pencil and paper or specifically designed software applications, have become increasingly prevalent in a variety of academic and corporate learning environments.

Due to space limitations, the following cited works only begins to hint at the extensive research base for cognitive mapping and the variety of techniques available for presenting knowledge structures.

Reviewing 20 studies involving the use of models for the facilitation of learning scientific material, Mayer (1989) concludes that the use of conceptual models as educational tools allow students more easily to assimilate knowledge through meaningful learning into valuable conceptual frameworks. Using models was also shown to greatly improve levels of conceptual retention and creative problem solving (in spite of a slight decline in verbatim retention of accompanying textual material).

In a comparison of concept mapping exercises with writing essays as learning tools, Markow & Lonning (1998) found no significant differences in post instructional achievement test scores between freshman chemistry students constructing concept maps and those writing essays. The authors argue that the multiple-choice achievement test is inadequate in probing conceptual understanding. They suggest further research to determine whether current concept-mapping methods provide an honest indication of student cognitive structure.

Level of knowledge is found to contribute to structural differences. Ayersman (1995), in a study of students' hypermedia knowledge, reported that low-knowledge students constructed hierarchical concept maps whereas high-knowledge students constructed complex maps in other formats such as web structure. He suggests that deliberate instruction is required to increase metacognitive accuracy and map complexity. Similar results were found in a research setting. Wang (1999) asked researchers to construct concept maps on 10 completed research topics. For each topic, a set of terms was extracted from the researcher's descriptions of the topic at both the outset and the completion of the project. The expert (the original researcher of the project), and another non-expert (also a researcher in the same subject field), mapped each topic respectively, resulting in two maps per topic. The non-experts tended to generate hierarchically structured maps while the experts produced complex structures such as web (radial), cross (two intersecting hierarchical structures), or three-dimensional configurations.

Effect of learning style on concept mapping was investigated by Reed & Oughton (1998) in a collaborative concept-mapping setting. Six groups consisting of various student combinations of similar or dissimilar domain knowledge and Kolb learning style constructed collaborative concept maps on the topic "hypermedia."

The groups consisting of high-knowledge students of a variety of learning styles were the most productive (Productivity was measured by number of nodes, links, and hierarchical complexity).

Learner affect has been a concern in education, Jegede et al. (1990) found "a tendency for the concept mapping strategy to significantly reduce anxiety toward biological achievement in males," possibly a result of the self-empowerment through learner cognizance and increased control of the assimilation process (p. 956).

Concept mapping has seen use primarily in "hard" science fields such as physics, chemistry, and biology. Todd & Kirk (1995) were the first to carry out a three-year study on the usage of concept mapping in one of information science foundation courses. Their students were carefully guided to learn and use concept mapping to draw important concepts and their conceptual relationships during the course. The majority of their students benefited from the experience and were positive about the usefulness of the tool. Negative experiences include initial anxiety about the process, maps' lack of details, or tediousness. They also observed changes in attitude along the process.

Various methods and techniques have been developed and evaluated for concept mapping. For example, Beatty & Gerace (2002) developed a series of computer based term-association tasks (ConMap) for educing the conceptual linkages between physics concepts, which served as a preliminary step in the development of metacognitive tools for mapping domain knowledge structures. Through a study observing 16 undergraduate physics students, the authors conclude that Term Prompted Term Entry (TPTE) tasks elicit a conceptual knowledge core, while hand drawn concept maps frame these core concepts within a more complete representation of student knowledge structure.

Czuchry and Danserau (1998), in a study of 108 undergraduate psychology students' retrieving personally relevant information, found that link-guided mapping and listing produced significantly fewer ideas (represented by nodes or listed terms) than with unguided mapping or listing. Attributing factors might include the increased cognitive load created by propositional links, time constraints, or limited experience with mapping and term association techniques.

Goldsmith & Johnson (1990) used matrices to assess domain knowledge structures quantitatively. Students were given 435 pairs of concepts (derived from 39 concepts) to rate relatedness using a 7-point scale. The resulting proximity matrices were compared to evaluate similarity. They found that agreement between students and the instructor increased over the semester; higher achievers had more similar conceptual structures among themselves than poorer students (Goldsmith & Johnson, & Acton, 1991).

While the usefulness of graphical representations of knowledge structure as an educational tool is widely recognized, there is controversy as to whether or not concept mapping is any more effective in evaluating students' levels of knowledge than traditional methods (such as standardized testing). Some researchers believe it is; others argue that the difficulties encountered in scoring concept maps limit their use to teaching and learning devices. Tsai and Huang (2002) note that while concept mapping allows for an assessment of student knowledge structure, authenticity issues are raised due to the possibility of bias introduced concerning the purpose and creation of concept maps during preliminary instruction.

In a series of rigorous studies to examine concept-mapping methods, Ruiz-Primo and associates reviewed different mapping techniques with various response formats and scoring methods. The wide variation in procedure and interpretation raises both validity and reliability issues in the use of any one method for accurately defining declarative knowledge structure (Ruiz-Primo & Shavelson, 1996). The following concept-mapping methods were carefully compared: (1) construct-a-map from scratch, in which concepts were provided; (2) fill-in-the-nodes, in which a skeleton map of missing nodes were provided; and (3) fill-in-the-lines, in which a skeleton map of missing lines were given. They concluded that these mapping techniques provided different pictures of domain knowledge structures; hence criteria for the determination of appropriate mapping techniques should be developed on the basis of sound cognitive theory and the specific domain knowledge under investigation. Techniques for soliciting knowledge representations should measure a well-defined cognitive activity. (Ruiz-Primo, Schultz, & Shavelson, 2001; Ruiz-Primo, et al., 2001)

Looking forward, research into the graphical representation of domain knowledge has important implications concerning IR system design, with the connection between cognitive structure and machine recognition or emulation of these structures (Lin, Marchionini, & Soergel, 1993). Schvaneveldt, Dearholt, & Durso (1988) explore applications of graph theory to the measuring of network structures (Similarities and differences can be compared quantitatively). Wang (2000) illustrates the conversion of concept maps into matrices: a network can be represented by the adjacency matrix and transformed to the distance matrix.

Among different knowledge structures, hierarchical organization has been the most-used method in IR. Both classifications and thesauri are important tools of IR systems. In a recent study (Bilal & Wang, in review), eleven seventh graders were given science concepts selected from two popular children's Web directories (Yahooligans! and KidsClick!) to construct hierarchical maps. They found that children were able to map science concepts hierarchically; maps of concrete concepts showed greater overlap among the children than those of abstract concepts; similarly, overlaps of conceptual structures between the children and the Web directories were greater for concrete concepts than abstract concepts. Children also tended to relate concepts situationally or experientially while the Web directories organize concepts in a disciplinary approach. These differences can hinder finding information via Web directories.

Lykke Neilsen (2002) proposed a novel approach to thesaurus construction. In her study, fifty domain subject experts were given a list of 100 words as stimuli to generate related words. The results were used to design an associative thesaurus. The associative thesaurus was further tested alongside a traditional literary-based thesaurus to compare user behaviors and search performances. Although no significant difference was observed, the method is definitely a "useful inspiration, whether compiler or searcher" (176).

While the impact of conceptual modeling on IR is fertile ground for exploration, researchers face epistemological, contextual, and semiotic challenges concerning incorporating human conceptual representations in IR systems.

Methodology

This study adopts a quasi-experimental method to observe the development of knowledge and cognitive structures over the natural span of a course. Classroom teaching and learning provides an appropriate setting for investigating knowledge structures in that learners share the same goals and assignments guided by an instructor with a well-developed syllabus. Regular meetings over the semester also enable interactions among the members of the class. We believe that certain classroom environments contribute to overlap in knowledge structures and data collection can be implemented in a structured manner during the semester.

Measurements & Tasks

Knowledge structure is measured in this study by size and content of vocabulary and relationships between concepts. For the latter, a Concept Matrix (Appendix A) is the mathematical representation of the knowledge structure, which can be transformed into a visual diagram—concept map, and vice versa.

Based on the theory of human memories (Best, 1995), human beings hold their knowledge in Long-Term Memory (LTM) and perform cognitive tasks using the activated knowledge available in Short-Term Memory (STM). Thus, two cognitive tasks are appropriate to measure knowledge structure: (1) listing terms representing concepts of the subject domain, a timed *recall* task; (2) identifying relationships between concepts, an untimed *recognition* task.

In order to perform the *recall* task, concepts must be <u>retrieved</u> from LTM into STM. Related concepts are <u>retrieved</u> via spreading activation. The speed of recalling a concept depends on the strength of the link between the concept and the previously activated concept. Therefore, the time lag between terms can indicate the strength of their semantic association. The initial cue that activates concepts in LTM comes from instruction given at the beginning of the session. For example, the experimenter says: "In the next [*x*] minutes, please write down the terms you know about [*course name*] in the order that they come to mind. Keep going until I tell you to stop or you run out of terms." Here the *course name* is the only initial stimulus or cue.

For the *recognition* task, all the concepts and relationship types are given; they serve as stimuli to activate conceptual structures in LTM. Our instrument focuses on capturing the hierarchical structures of a learner's domain knowledge.

Instruments

The instruments for data collection include term listing, Concept Pairwise, concept mapping, freewriting, structured diary, and the Kolb Learning inventory.

Two versions of term listing were used: (1) entering terms into an Excel sheet during a timed session and (2) a Web-based form (Appendix E), which times the session while adding a timestamp to each term upon entry.

Concept Pairwise was designed to collect conceptual relationships. The participants were given a list of concept pairs (Appendices B) and a list of relationships (Appendix C). For each pair, the subject assigned a relationship type and a number representing associative strength. The Concept Pairwise can be easily transformed into a Concept Matrix for analysis (Appendix D). Two versions were used: (1) paper-pencil format and (2) Web version (Appendix F). In comparison, the Web version has improved usability, but requires access to a computer and the Web. Strength of term association rating was omitted for the Web version because the timelags between term entry in the first session were recorded as an indicator of concept association strength.

Software programs such as Inspiration exist for concept mapping to produce graphic maps with nodes for concepts and lines (with or without arrows) for relationships. Depending upon the situation, it is often easier to construct a concept map using paper and pencil to avoid learning additional computer skills. We use simply a large sheet of paper and blank Post-it notes. The participants wrote each concept on a note and place the notes in a meaningful configuration on the sheet.

A timed, in-class freewriting exercise on *what the writer knows about the subject* is a complementary instrument and is much easier than listing terms, Concept Pairwise, or concept mapping. Although the data are less structured and difficult to apply to the concept matrix, careful content analysis can provide useful triangulation.

Participants were asked to keep diaries focusing on aspects of research methods under predefined headings outside class document thoughts and experiences. Collected regularly, these diaries provide important feedback to the instructor during the course and data collection period so that issues are immediately attended and solved. For example, a few students felt nervous due to the noise of neighbors rapid keyboarding during the session.

The Kolb Learning Style Inventory is a standard psychological instrument for measuring individual differences (Jonassen & Grabowski, 1993). Learners are classified to one of four types: accommodator, assimilator, converger, or diverger. Accommodators prefer "hands-on" experience and tend to put ideas into action and practice relying on "gut" feelings rather than logical analysis. In contrast, assimilators prefer putting information into a concise and logical framework or structure and tend to think things through. Convergers enjoy relating theory to real world situations and tend to be technical. Divergers prefer to look at a situation from multiple perspectives and tend to be imaginative and open-minded. For the task of organizing concepts, the assimilating learning style has a advantage while both convergers and accommodators may rely on experiences to build conceptual relationships using a situational approach. When lacking experiences, convergers and accommodators may encounter difficulty representing. knowledge structures. Divergers may undergo more dynamic changes in knowledge structure during the learning process than other styles of learners.

Procedure

At the beginning of the course, (1) students were introduced to the study and invited to participate; anonymity and confidentiality measures were explained. (2) Each participating student randomly picked a data identification number from a box. (3) All students took the Kolb inventory. (4) A fifteen-minute pre-course freewriting exercise was given in class on "what I know about [*course title*]." (5) A trial experimental session was administered to practice using the instruments.

Knowledge structures were surveyed three times during the semester: (1) the first data collection was performed after completion of one third of the semester; (2) the second after two thirds of the semester; (3) the final data collection at the end of the semester.

In addition, participants kept structured diaries on a weekly basis. The predefined heading relating to knowledge structure is Experiences and Thoughts as a Research Participant (Experiences and Thoughts on the Knowledge Survey). As an measure of anonymity, diaries were not linked to individual students during data analysis, but provided triangular data as a whole and served as a feedback channel. Diaries had no baring on students' course grade

At the end of the course, the participants constructed a concept map using Post-it notes on a large blank sheet. They were also given an in-class post-course freewriting exercise on "what I know about [*course title*]."

Methodological Notes

The anonymity and confidentiality measures guarantee that data are not linked to individual students, therefore a written consent is not required by research compliance. For the second study, concept mapping was adopted as a learning and evaluation device. Students were offered the option to choose either a traditional test or concept mapping for the 25% of the course grade. Anonymity is achieved by replacing student identity with a random number for data for analysis at the end of the course.

Both the procedure and instruments were revised several times during the pilot tests and the actual experiment runs. The original design for listing terms asked the participants to stop when readily running out of terms. Although some students stopped in a few minutes others continued in excess 20 minutes. This problem was corrected by imposing a time limit.

The original design for Concept Pairwise used the participant's own terms from task one to generate pairs. However, neither the size of the lists nor the quality of the listed terms warranted meaningful results. If a participant had listed too many terms in task one, his/her Concept Pairwise for task two became unmanageable. The design was revised to provide a standard Concept Pairwise, in which the instructor selected terms based on teaching materials and the results of previous sessions. As a result of these changes, the first data collection was treated as a trial session and excluded from most analyses.

At the end of the first study, we began use of a Web interface to implement instruments for term listing and identifying conceptual relationships. (Appendices E and F).

Study One

This study explores the first five research questions. Eleven students, enrolled in a summer research methods course, voluntarily participated in the experiment to earn bonus course points (the experimenter stamped each session on a participation form). The 3-credit course was scheduled for 4.5 weeks; the class met for 3.5-hours on Monday, Wednesday, and Friday. The subjects were observed three times completing two tasks: listing concepts in Excel and identifying relationships using Concept Pairwise. All three experimental sessions were scheduled on Mondays as the first item on the agenda. Each data session lasted about one hour. A research assistant administered the sessions without the instructor being present.

At the beginning of the course, the students completed the Kolb inventory and participated in a freewriting exercise about what they "know about *research methods*." During the course, the students wrote about their experience as a research participant in their weekly diary entries. At the end of the course, each student constructed a visual concept map using Post-it notes to place concepts on a large sheet of paper.

The participants were assured anonymity, which was implemented by having each participant draw a folded paper containing a unique number data identification. Diaries and freewriting about the subject are not linked to the knowledge structure data.

Preliminary results

The experiment resulted in both quantitative and qualitative data. For quantitative data, statistical analysis focuses on the descriptive features of vocabulary and conceptual structures across time; differences in means and possible associations are also explored. For qualitative data, content analysis focuses on comparison of post and pre knowledge captured in the texts of freewriting as well as the Concept Matrices between the last two sessions.

Table 1 indicates that the size of vocabulary measured by both the number of terms and the number of concepts (after resolving synonyms and word variations) increased over time. The increase between session 2 and session 3 is significant at the 0.001 level using paired t-test. The correlation between the number of concepts in the two sessions is also significant (p= 0.001). Conceptual overlapping across participants (inter-overlapping) ranges from 1% to 44%; conceptual overlapping across sessions (intra-overlapping) ranges from 5% to 23%.

Table 1. Descriptive Statistics

Session	Min	Max	Mean	SD
Number of Terms Listed				
1*	9	45	17.64	12.49
2	9	40	22.36	11.45
3	16	52	28.64	12.00
Number of Concepts Listed				
2	8	39	21.36	10.64
3	16	52	28.45	11.81
Number of Sentences in Free-writing				
Pre	7	15	11.55	2.30
Post	5	26	14.45	7.29

* The first session was not timed.

In addition to the given 15 types of conceptual relationships, the participants defined four more types of relationships (Appendix C). We deliberately eliminated the *related* relationship in order to aspire to specified relatedness between two concepts, such as *Cause-effect*, *Temporal*, and *Spatial*. Two subjects added *Related* (RT) as a type, which seems to derive from their familiarity with thesauri structures. Two subjects added XY for remotely related and NR for not related. The individual adding NR obviously overlooked that the relationship was already on the list with the code X; this subject realized the oversight in the later sessions. We found that the mean for relationship types was 6.5, ranging between 5 and 10. All of the relationship types except one (*Spatial*) were used; the most used relationship types were *Narrow* (NT) and *Broad* (BT) by nine subjects.

For each concept pair, the assigned relationship types show a diversity from 3 to as many as 9 types. For example, the concept pair (*data collection* and *data analysis*) was assigned 7 relationship types with the most used relationship being *Temporal* (by 3 subjects in Session 2 and 4 subjects in Session 3). The lowest inter-consistency was 18%; that is, at most there were two students that agreed on a relationship type. However, there also exist high agreements concerning specific relationship type across subjects. The concept pair (*quantitative approach* and *qualitative approach*) was assigned *Antonym* or *Not Related* by 9 subjects in both sessions—82% agreement. The average inter-agreement for all the pairs is 39% in session 2 and 36% in session 3.

The intra-consistency (the overlapping of the relationship type for each pair in two experimental sessions) ranges from 5% to 90%, with the average of 45%.

We grouped the 15 relationship types into six categories: (1) merging *Class-Member* and *Whole-Part* with *Broad-Narrow*; (2) *Member-Class* and *Part-Whole* with *Narrow-Broad*; (3) *Not Related* with *Remotely Related*; (4) keeping *Synonym*; (5) *Antonym*; (6) merging all the rest as RT. After the grouping, the intra-consistency is between 33% and 90% with the average of 57%.

The eleven subjects have four learning styles: six assimilators, two convergers, two convergers and one accommodator. Due to the uneven number in each group, we can only speculate on group differences. The means of relevant concepts recalled and standard deviations indicate that differences are smaller within each group than the sample as a whole. For example, the largest group of assimilators (N=6) recalled 17 (SD 6.39) and 23 (SD 5.98) concepts in the two sessions. In comparison, the sample's SDs for both sessions are almost twice of the group's SDs. The group of two divergers recalled the most concepts. In the task of identifying conceptual relationships, we found that the assimilators were the most diverse in intra-consistency.

The post-freewriting exercise resulted in substantially more relevant concepts than the pre-freewriting exercise; changing from knowledge about *research* in general to knowledge about *research method*. The final visual maps are diverse in configuration and hard to compare without concurrent verbal data to explain why a term was placed in a certain way. It is not possible to interview the subjects about the map because of the anonymous data requirements. However, data is useful when related to the Concept Matrix to verify consistencies.

Reflected in the subjects' diary entries, incorporating a real experiment into a research method course is definitely a good practice. After initial unfamiliarity with the instruments, subjects quickly felt comfortable about the tasks of listing concepts and identifying relationships. The need to revise the instrument and procedure also provided students with a real-life case about conducting research. Students like the idea of concept mapping to enhance concepts during a regular semester. they often felt terms remained on the tip of their tongues during the task of listing.

Study Two

The second study was conducted during a regular semester in a required introductory course on organization of information. This study explores concept mapping as a metacognitive tool to help learners build correct conceptual structures and as an alternative to traditional quizzes and tests in introductory courses. The purpose of this study is to focus on the last two research questions, but also address the first four questions. The two tasks, listing terms and identifying relationships, were titled "knowledge surveys" and use a Web interface (Appendices E & F). Students can choose to take either knowledge surveys or exams to complete 25% of the course assignments. All students (N=34) chose knowledge surveys.

The results from Study One suggest that average learners handle about 6 relationship types and the most-used type is the super- and sub-ordinate relationship, hierarchical in nature. We selected the following five categories for this study: (1) BT, $term_i$ <is a broader concept than> $term_j$; (2) NT, $term_i$ <is a narrower concept than> $term_j$; (3) ST, $term_i$ <is a synonym of> $term_j$; (4) RT, $term_i$ <is a related concept of> $term_j$; (5) NR, $term_i$ <is not related to> $term_j$. The Web form provides a drop-down list of the five relationships in natural language (The two-letter codes were not displayed). These relationships were defined in a guide available on the course Webpage during the semester. For example,

- Broad-narrow relationships include
  generic-specific
  (e.g., OPAC <broader than> UTK online catalog)
  whole-part
  (e.g., OCLC FirstSearch <broader than> WorldCat)
  class-member
  (e.g., IT standards <broader than> ASCII)
- Not related terms may be used for remotely related or antonym.

In actuality, all given pairs did not have antonyms.

As an in-class exercise, students worked in groups to draw visual maps. Before this exercise, the instructor reviewed an earlier figure in which information retrieval and access consists of four interrelated main components: (1) information objects (documents), (2) information technology (standards, systems, and algorithms), (3) user, and (4) interface. Each group picked two of the components to for the purpose of constructing concept maps; one of the maps must be selected from the first two components. The results were shared in class.

Two weeks before the final knowledge survey, the instructor overviewed major concepts delivered in the course and provided a concept map of the hierarchical organization of the major concepts. The map starts from the main subdivisions, the same four main components as above (Appendix G).

Preliminary results

Similar to the first study, students could not focus on important concepts early in the semester. Many students listed general terms from the vocabulary of the field rather than subject matter terminology of the subject matters relating to the course. Terms like *library*, *information* and *data* were listed. In a diagnostic session following the trial data collection, students were introduced to the differences between "terminology as technical terms defined/learned in this course" and "vocabulary of daily life or the field"; and the importance of focusing on "terms likely to occur in a glossary of a book on organization and representation of information, metadata for digital libraries" and "quality not quantity in listing concepts."

The number of listed terms ranges from 16 to 98 with an average of 49 in session one (10-minute) and 20 to 78 with the average of 42 in session two (7-minute). The most-shared terms are the first 20 terms: ten students shared 11 terms in session one and ten students shared 14 terms in session two. More than half of the class listed the following terms in the last two knowledge surveys: *MARC*, *abstract*, *DDC*, *AACR2R*, *Access points*, *metadata*, *Dublin Core*, *authority control classification*, *hierarchical structure*, *indexing*, *controlled vocabulary*, *ISBD*, *call number*.

Timelags between terms show patterns of pauses (a pause is defined as the timelag greater than twice of the average timelag). Pause signal the boundaries of clusters. A cluster consists of closely associated concepts. For example, one student entered 37 terms with three major pauses at 12th, 26th, and 34th terms. The four clusters are as follows in the order of the terms being entered:

1. *cataloging, catalogers, authority control, standard numbers, ISSN, ISBN, ISBD, AACR2R, MARC, USMARC, UNIMARC*
2. *Z39.50, ANSI; search algorithm, abstract, indicative-informative abstract, informative abstract, indicative abstract, structured abstract, extended abstract, annotation, extract, card catalog, OPAC, access points*
3. *LC catalog number, cutter number, Dewey Decimal, open stacks, closed stacks, cutter rules, classification schedules, LC Subject Headings*
4. *tags, fields, colon classification, metadata*

The above example illustrates that the first cluster includes concepts related to cataloging and standards, the second cluster has concepts related to searching, indexes and abstracts, the third and fourth clusters are concepts related to records and subject access devices. It is not surprising to see certain concepts are strongly associated: AACR2R with MARC and classification with subject headings (See the last two terms in 3rd cluster and Figure 1)

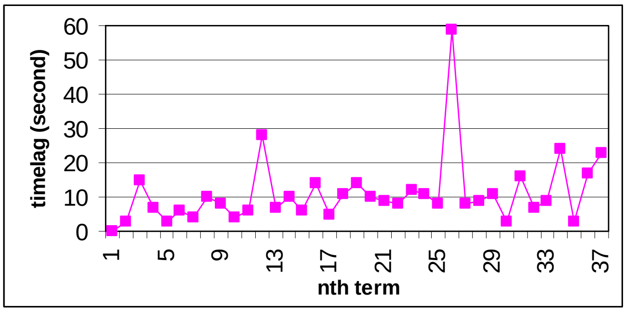

Figure 1. Pausing patterns

For the second task, identifying relationships between concepts, we found that students' organizational structures of the concepts were less diverse towards the end of the semester and inter-agreements improved significantly. For example, *authority control* was correctly recognized as a broader concept of *name headings* by 73% in the last session but by only 45% in the previous session.

The intra-agreements range from 9% to 81% with the average of 47%. Further analysis is needed to verify if lower intra-agreement scores indicate greater improvements in knowledge structures.

Because the knowledge surveys were used in lieu of traditional tests, the instructor adopted an evaluation system. For term listing, major terms score two points, acceptable terms score one point and zero points for trivial terms (not in the terminology of the course). For identifying relationships, correct relationships score two points, acceptable relationships score one point, and zero points for misconceptions. For the above example, the pair *authority control* and *name headings*, two points were given for selecting <broader than>, one point for <related> or <synonym>, zero point for <not related>. The raw scores are normalized and converted to a final score toward the 25% grade.

Students commented that mapping concepts hierarchically helped them to put concepts into a structured framework to see the big picture of the course without getting lost in minute details. Knowledge surveys made them think and focus on the most important concepts and how these concepts relate. It was useful to differentiate *terminology* from *vocabulary*. In addition, a knowledge survey relieves test anxiety and pressure before tests. However, similar to Todd and Kirk's results (1995), a few students were uncomfortable about the new tool and wanted a "correct" map from the instructor.

Summary and Conclusions

We developed and tested two instruments to collect data on learners' knowledge structures during two introductory courses. Instrument design was based on the theory of long-term and short-term memories responsible for the cognitive tasks involved: *recall* and *recognition*. The *recall* task retrieves terminology of domain knowledge and the *recognition* task identifies relationships between given pairs of concepts.

Our preliminary analyses show some interesting results pertaining to our research questions. Students' vocabularies increased over the course; concepts became more accessible from five terms per minute to six terms per minute. The quality of the vocabulary also increased as the learner became more focused on terminology. The most used relationship type was hierarchical, followed by related and temporal relationships in the research methods course. Overlapping in concepts and inter-agreement in relationship types increased towards the end of the semester. Intra-agreements were diverse; one assumption is that the lower the intra-agreements, the greater the improvements in knowledge structures. This assumption will be tested in further analysis. The students with the same learning style produced similar term lists.

A tentative conclusion is that the knowledge survey using the two instruments is a potentially useful device to measure domain knowledge for introductory courses (in which the purpose is to learn basic concepts and build a foundational knowledge structure for advanced studies).

Further analysis will focus on integrating different data to gain insight into the development of knowledge structures, the validity of measurements, and the improvement of instruments. The findings from this line of research have implications for both education and IR: concept mapping is a valuable metacognitive tool in teaching introductory courses in information sciences and the methodology has the potential to capture conceptual structures and incorporate them into IR systems to support differences in searchers' knowledge levels.

Notes

[1] Inspiration, a computer tool for developing ideas and organizing thinking, defines a *concept map* as a hierarchical diagram to represent a set of concepts beginning with the most general or important and working down to more specific.

[2] Inspiration defines a *web* or an *idea map* as a structure that places the main idea (core concept) or the problem at the center expanding to related ideas and concepts.

REFERENCES

Ayersman, D.J. (1995). Effects of knowledge representation format and hypermedia instruction on metacognitive accuracy. Computers in Human Behavior, 11 (3-4), 533-555.

Ausubel, D. P. (1963). The psychology of meaningful verbal learning. NY: Grune and Stratton.

Ausubel, D.P., Novak, J.D., & Hanesian, H. (1978). Educational psychology: A cognitive view. New York, NY: Holt, Rinehart and Winston.

Beatty, I.D., & Gerace, W.J. (2002). Probing physics student's conceptual knowledge structures through term association. American Journal of Physics, *7*0(7), 750-758.

Best, J. B., (1995). Cognitive Psychology (4th. ed.). St. Paul: West Publishing Co.

Bilal, D., & Wang, P. (in review). A study of children's conceptual structures of science categories and the design of web directories. [submitted for publication on May 5, 2004]

Czuchry, M., & Dansereau, D.F. (1998). The generation and recall of personally relevant information. Journal of Experimental Education, 66(4), 293-315.

Goldsmith, T.E. & Johnson, P.J. (1990). A structural assessment of classroom learning. In W. Schvaneveldt, (Ed.), Pathfinder associative networks: Studies in knowledge organization (pp. 241-254). Norwood, NJ: Ablex Publishing Corporation.

Goldsmith, T.E., Johnson, P.J., & Acton, W. H. (1991). Assessing structural knowledge. Journal of Educational Psychology, 83 (1), 88-96.

Ingwersen, P (1992). Information Retrieval Interaction. London: Taylor Graham.

Jegede, O.J., Alaiyemola, F.F., & Okebukola, P.A.O. (1990). The effects of concept mapping on student's anxiety and achievement in biology. Journal of Research in Science Teaching, 27(10), 951-960.

Jonassen, D.H., & Grabowski, B.L. (1993). Handbook of Individual Differences: Learning and Instruction. Hillsdale, NJ: Lawrence Earbaum Associates.

Jonassen, D. (2000). Knowledge is complex: Accommodating human ways of knowing. In D. Soergel, P. Srinivasan, & B. Kwasnik (Eds.), Proceedings of the 11th ASIS&T SIG/CR Classification Research Workshop (pp. 1-7), Silver Spring, Maryland: American Society for Information Science.

Lakoff, G. (1987). Woemn, Fire, and Dangerous Things: What Categories Reveal about the Mind. Chicago, IL: The University of Chicago Press.

Lin, X., Marchionini, G., & Soergel, G. (1993). Category-based and association-based map display by human subjects. Proceedings of the 4th ASIS SIG/IR Classification Research Workshop, October (pp. 147-164). Columbus, OH.

Lykke Neilsen, M. (2002). The word association method: A gateway to work-task based retrieval. Dissertation. Åbo, FI: Åbo Akademi University Press.

Markow, P.G. , & Lonning, R.A. (1998). Usefulness of concept maps in college chemistry laboratories: Students' perceptions and effects on achievement. Journal of Research in Science Teaching, 35(9), 1015-1029.

Mayer, R.E. (1989). Models for Understanding. Review of Educational Research, 59(1), 43-64.

Moody, J.W., Will, R.P., & Blanton, J.E. (1996). Enhancing knowledge elicitation using cognitive interview. Expert Systems With Applications, 10(1), 127-133.

Novak, J.D. (1998). Learning, creating, and using knowledge. Mahwah, NJ: Lawrence Erlbaum.

Novak, J.D. (1990). Concept maps and Vee diagrams: Two metacognitive tools for science and mathematics education. Instructional Science, 19, 29-52.

Novak, J.D., & Gowin, D.B. (1984). Learning how to learn. Cambridge, UK: Cambridge University Press.

Reed, W.M., & Oughton, J.M. (1998). The effects of hypermedia knowledge and learning style on the construction of group concept maps. Computers in Human Behavior, 14(1), 1-22.

Ruiz-Primo, M.A., Schultz, S.E., Li, M., & Shavelson, R.J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. Journal of Research in Science Teaching, 38(2), 260-278.

Ruiz-Primo, M.A., & Shavelson, R.J. (1996). Problems and issues in the use of concept maps in science assessment. Journal of Research in Science Teaching, 33(6), 569-600.

Ruiz-Primo, M.A., Shavelson, R.J., Li, M. & Schultz, S.E. (2001). On the validity of cognitive interpretations of scores from alternative concept-mapping techniques. Educational Assessment, 7(2), 99-141.

Rumelhart, D.E., & Norman, D.A. (1985). Representation of knowledge. In A.M. Aitkenhead, & J.M. Slack (Eds.), Issues in cognitive modeling (pp. 15-62), Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Schvaneveldt, R.W., Dearholt, D.W., & Durso, F.T. (1988). Graph theoretic foundations of pathfinder networks. Computers and mathematics with applications, 15(4), 337-345.

Todd, R. & Kirk, J. (1995). Concept mapping in information science. Education for Information, 13(4), 333-47.

Tsai, C,-C, & Huang, C,-C. (2002). Exploring students' cognitive structures in learning science: A review of relative methods. Journal of Biological Education, 36(4), 163-170.

Wandersee, J.H. (1990). Concept mapping and the cartography of cognition. Journal of Research in Science Teaching, 27(10), 923-36.

Wang, P. (1999). An empirical study of knowledge structures of research topics. In L. Woods (Ed.), Proceedings of the 62nd ASIS annual meeting (pp. 557-586), Medford, NJ: Information Today.

Wang, P. (2000). Comparing Cognitive Maps Using Graph Algorithms**.** In D. Soergel, P. Srinivasan, & B. Kwasnik (Eds.), Proceedings of the 11th ASIS&T SIG/CR Classification Research Workshop (pp. ???), Silver Spring, Maryland: American Society for Information Science.

Appendix A. Concept Matrix

R represents the type of relationship between concepts;

$\mathbf{R}_{ij}$ defines a cell in a Matrix at the intersection of $\mathbf{T}_i$ and $\mathbf{T}_j$

T represents term/concept;

$\mathbf{T}_i$ defines a *row*-term

$\mathbf{T}_j$ defines a *column*-term

e.g., $R_{ij}(t_i, t_j)$ = NT if T_i is narrower than T_j

Similarly,

A represents the strength of the association between concepts;

$\mathbf{A}_{ij}$ defines a cell in an adjacency matrix at the intersection of $\mathbf{T}_i$ and $\mathbf{T}_j$

e.g., $A_{ij}(t_i, t_j) = 5$ is the strength of T_i and T_j

Appendix B. Concept Pairwise

Term $_I$	Term $_j$	Relationship	Association
data collection	data analysis		
experiment	data collection		
instrument	experiment		
interview	instrument		
qualitative	interview		
quantitative	qualitative		
experiment	data analysis		
instrument	data collection		
interview	experiment		
qualitative	instrument		
quantitative	interview		
instrument	data analysis		
interview	data collection		
qualitative	experiment		
quantitative	instrument		
interview	data analysis		
qualitative	data collection		
quantitative	experiment		
qualitative	data analysis		
quantitative	data collection		
quantitative	data analysis		

Appendix C. Relationship Coding List

Codes	Meaning	Notes
SN	first term is *synonym of* second term	
AN	first term is *antonym of* second term	
X	not related	
NT	first term is *narrower* than second term	All these relationship types define hierarchical relationships. Two terms are related as super- and subordinate concepts or vice versa
BT	first term is *broader* than second term	
PW	*part-whole* relationship such as wheel-Car	
WP	*whole-part* relationship such as hand-finger	
CM	*class-member* relationship e.g. chair-wheel chair	
MC	*member-class* relationship e.g. chair-furniture	
CE	*cause-effect* e.g. SARS cause fever	All these relationship types define associated relationships
EC	*effect-cause* e.g. headache is caused by stress	
UF	*used for* e.g., lawnmower used for cutting grass	
UB	*used by* e.g., hammer used by handyman	
SP	*spatial* e.g., monkeys live in mountain	
TM	*temporal* e.g., before, after, simultaneous, etc.	
Defined by the subjects		
UI	*Use in*	Added by participant
RT	Related terms	Used heavily by one person
XY	Remotely related	
RF	Required for	*instrument* RF *experiment*

Appendix D. Example Concept Matrix

	data analysis	*data collection*	*experiment*	*instrument*	*interview*	*qualitative*
Data collection	TM					
Experiment	TM	NT				
Instrument	X	UF	UF			
Interview	TM	NT	X	~~UF~~		
Qualitative	BT	BT	X	X	~~UF~~	
Quantitative	BT	BT	BT	X	X	AN

Note: the two crossed cells illustrate that participants frequently transpose the pair.

Appendix E. Simple interface for listing

http://web.utk.edu/~peilingw/ks/screenshots.doc

Appendix F. Simple interface for identifying relationship

http://web.utk.edu/~peilingw/ks/screenshots.doc

Appendix G. Hierarchical organization of important concepts for the course

http://web.utk.edu/~peilingw/ks/IS520map.doc