Latent Semantic Analysis Captures Causal, Goal-oriented, and Taxonomic Structures

Arthur Graesser (a-graesser@memphis.edu)
Department of Psychology, University of Memphis, CAMPUS BOX 526400
Memphis, TN 38152 USA

Ashish Karnavat (akarnavat@hotmail.com)
Department of Computer Science, University of Memphis, CAMPUS BOX 526429
Memphis, TN 38152 USA

Victoria Pomeroy (vpomeroy@memphis.edu)
Department of Psychology, University of Memphis, CAMPUS BOX 526400
Memphis, TN 38152 USA

Katja Wiemer-Hastings (kwiemer@cc.memphis.edu)
Department of Psychology, University of Memphis, CAMPUS BOX 526400
Memphis, TN 38152 USA

and the Tutoring Research Group

Abstract

Latent Semantic Analysis (LSA) has been used to represent the domain of computer literacy in AutoTutor, a fully automated computer tutor. The analyses in the present study support the claim that the 200-dimensional LSA space captures aspects of the structured mental models that underlie computer literacy. Knowledge structures were constructed that contained causal networks, goal/plan/action hierarchies, and taxonomic hierarchies. The proximity of a pair of nodes (i.e., concept, state, event, action, goal) in these structures predicted the cosine similarity scores that are routinely computed in LSA analyses.

Representing World Knowledge with Conceptual Graph Structures

World knowledge has traditionally been captured by knowledge structures throughout the history of cognitive science, artificial intelligence, and discourse processes. The knowledge structure structures include semantic networks, taxonomies, causal networks, planning networks, ontological trees, spatial region hierarchies, and various other classes of conceptual graph structures (Golden, 1997; Graesser & Clark, 1985; Kiel, 1979; Lehmann, 1992; Lenat, 1995; Norman & Rumelhart, 1975; Schank & Abelson, 1977; Trabasso, van den Broek, & Suh, 1989; Sowa, 1983). A knowledge structure contains a set of categorized nodes that refer to concepts, events, processes, states, actions, goals, and other ontological classes. The nodes are connected by relational arcs that also are assigned to various categories, e.g., is-a, has-as-parts, cause, reason, enables, contains, etc. A particular package of knowledge may incorporate spatial composition, causal networks, goal hierarchies, taxonomic hierarchies and other viewpoints. All of these viewpoints allegedly can be represented as a set of categorized nodes that are integrated by a set of directed, relational arcs.

It is a time consuming, methodical task to map out knowledge structures for a domain of knowledge. Developers of expert systems and other knowledge based systems would require a decade to perform the knowledge engineering that is needed for a system of reasonable scope with widespread practical applications (Lenat, 1995). There are authoring tools that guide either experts or novices in the building of the knowledge structures (Williams, Hultman, & Graesser, 1998). The structures are built in a principled fashion that caters to the constraints of the composition rules, so guidance is needed to prevent illegal compositional structures. All of this takes training and experience that can be measured in months or years. However, conceptual graph structures are powerful theoretical entities because they support the intelligent procedures and processes that operate on the representations, as in the case of retrieval, classification, summarization, problem solving, question asking, question answering, and so forth.

The distance between two nodes in a conceptual graph structure is frequently regarded as a metric of conceptual relatedness. That is, the conceptual relatedness between nodes A and B decreases as a function of the number of arcs that exist on a legal path between A and B. For example, if 1 arc separates A and B on a causal chain, then A and B are strongly related, compared to the case where 4 arcs separate two nodes on a causal chain. The structural proximity between any two nodes that are connected by a legal path of arcs is designated as its structural-proximity (A, B).

Representing World Knowledge with Latent Semantic Analysis

Researchers have more recently turned to Latent Semantic Analysis (LSA) because it provides an approximation of the representation of world knowledge, but in a very short period of time -- measured in weeks, days or even hours. LSA is a statistical representation of a body of world knowledge that is reflected in a large corpus of textual documents (Landauer & Dumais, 1997; Landauer, Foltz, & Latham, 1998). LSA capitalizes on the fact that particular words appear in particular texts (called "documents"); the cooccurrence of words in documents reflects the constraints that exist in world knowledge. The input to LSA is a cooccurrence matrix that specifies the number of times that word W_i occurs in document D_j. These frequencies are adjusted with a logarithm transformation that also corrects for the base rates of words appearing across documents. A word is a distinctive index for a document to the extent that its occurrence in the document is above the base rate for that word across documents. A standard statistical method, called singular value decomposition, reduces the large WxD cooccurrence matrix to K dimensions (typically, 100 to 500 dimensions). Each word, sentence, or text ends up being represented as a weighted vector on the K dimensions.

The similarity or conceptual relatedness between two bags of words (A and B) is computed as a geometric cosine (or dot product) between the two vectors. The values normally range from 0 to 1. This LSA match between two language strings is designated as its LSA-match (A, B). The LSA match can be high even though there are few, if any words in common between the two strings. LSA allegedly goes well beyond simple string matches because the meaning of a language string is partly determined by the company (other words) that each word keeps (Landauer & Dumais, 1997).

The empirical success of LSA has been promising and sometimes remarkable. Landauer and Dumais (1997) created an LSA representation with 300 dimensions from 4.6 million words that appeared in 30,473 articles in Grolier's Academic American Encyclopedia. They submitted to the LSA representation the synonym portion of the TOEFL test, a test developed by the Educational Testing Service to assess how well non-native English speakers have mastered the words in the English language. The test has a four-alternative, forced choice format, so there is a 25% chance of answering the questions correctly. The LSA model selected the alternative that had the highest match with a comparison word. The LSA model answered 64.4% of the questions correctly, which is essentially equivalent to the 64.5% performance for college students from non-English speaking countries. LSA has had remarkable success in capturing the world knowledge that is needed to grade essays of students (Foltz, 1996), to assign texts to students of varying abilities to optimize learning (Wolfe, Schreiner, Rehder, Laham, Foltz, Kintsch, & Landauer, 1998), and to provide effective feedback in the training of summarization skills (E. Kintsch, W. Kintsch, Laham, Landauer, DePaula, Schreiner, Stahl, & Steinhart, 2000). There are now LSA-based graders of essays that assign grades to essays with the validity and reliability of human experts in composition (Foltz, 1996). In our research on computer literacy, LSA has been quite successful in evaluating the quality of college students' answers to deep reasoning questions and to the contributions of learners during the tutorial interactions with AutoTutor (Graesser, Wiemer-Hastings, Wiemer-Hastings, Harter, Person, & the TRG, 2000; Wiemer-Hastings, Wiemer-Hastings, Graesser, and the TRG, 1999).

The success of LSA is quite remarkable given that it was never designed to capture many of the traditional problems in language understanding systems, such as word order, syntax, quantification, and negation. There are other corpus-based probabilistic models that capture word order and syntax (Burgess, Livesay, & Lund, 1998; Charniak, 1993) but the present study focuses on the capabilities of LSA.

At this point, there is a great deal of uncertainty about what is being represented in the K-dimensional spaces of LSA. One optimistic possibility is that the K dimensions reflect ontological categories, semantic features, and structural compositions of mental models that would be directly adopted in structural theories of world knowledge representation. For example, a simple and straightforward assumption would be that particular banks of the K dimensions of LSA would have a one-to-one or many-to-one mapping onto ontological categories (Chi, Slotta, & de Leeuw, 1994; Keil, 1979), to conceptual primitives (Miller & Johnson-Laird, 1976; Norman & Rumelhart, 1975; Schank & Abelson, 1977), or to the domain-specific features that are associated with a particular topic. Very few researchers would go out on the limb and propose an elegant mapping between the K dimensions of LSA and sophisticated theories of world knowledge. However, most researchers would seriously entertain the possibility of weaker correspondences. At the other end of the continuum, there are researchers who believe that the K dimensions have nearly an arbitrary mapping to the attributes of mature theories of world knowledge (Landauer & Dumais, 1997).

A somewhat different question addresses whether the LSA space is capable of recovering aspects of the deeper mental models that underlie text (Forbus, Gentner, & Law, 1995), or what is sometimes called situation models (Kintsch, 1998). Foltz, Britt, and Perfetti (1996) reported evidence that suggested that LSA does capture mental model representations to some extent, whereas Perfetti (1998) has expressed doubts that LSA captures the representations and processes of psychological models. LSA may capture shallow knowledge rather than deep knowledge. That is, LSA may capture the sort of word associations that are reflected in the archives of dictionaries and encyclopedias, but may not penetrate the deeper mental models. On the other hand, LSA may be successful in capturing aspects of the deeper situation model. An accomplished expert on some topic certainly does know how to use the right bags of words at the right time; the systematic use of words in particular documents may be recovered in the LSA solution spaces. At this point in the science, however, there is not enough empirical evidence to support one position or another.

The present study hopes to shed additional light on what is captured by the LSA representations. An LSA space has been developed in the domain of computer literacy. This LSA representation has been used in a fully automated computer tutor, called AutoTutor (Graesser, Franklin, Wiemer-Hastings, & the TRG, 1998; Graesser et al., in press; Graesser et al., 2000; Wiemer-Hastings, Graesser, Harter, & the TRG, 1998). In addition to the LSA space, AutoTutor has dozens of conceptual graph structures that capture knowledge in a more structured form. The present study examines whether the structural composition of the conceptual graph structures can predict the LSA match scores. That is, is there a significant correlation between structural proximity and LSA match scores when we examine taxonomic hierarchies, causal networks, and goal structures? A positive correlation would support the claim that LSA spaces to some extent recover aspects of the mental models. A zero correlation supports the claim that the K dimensional LSA spaces have an unsystematic mapping onto structural theories of knowledge representation.

Corpus of Texts and LSA Space on Computer Literacy

A 200-dimensional LSA space was developed for the domain of computer literacy during the development of AutoTutor. The corpus of included (a) two books on computer literacy, (b) 30 articles that focus on hardware, operating systems, and the internet, and (c) AutoTutor's curriculum script of lessons, example problems + solutions, and questions + answers. An LSA analysis requires the preparation of a document by word (D x W) co-occurrence matrix. Each cell in the matrix specifies the number of occurrences of word W_i in Document D_j. In order to prepare the DxW matrix, the researcher needs to define what constitutes a document unit. A single document was defined as (a) a paragraph in the case of the textbooks and 30 articles and (b) a sentence that conveys a lesson, a good answer, or piece of a solution in the case of the curriculum script. An LSA analysis was performed on the 2.3 MB corpus of documents, yielding a solution with 200 dimensions.

The 200-dimensional LSA was validated in our assessments of AutoTutor (Graesser et al., 2000; Wiemer-Hastings et al., 1999). For example, Wiemer-Hastings et al. (1999) analyzed how well the LSA space on computer literacy could accurately evaluate a sample of 192 answers to the questions in the curriculum script. College students enrolled in the computer literacy course answered the questions in the curriculum script by typing in their answers into a web cite facility. The data were collected after the college students had read the relevant chapters in the book and had received a lecture on each macrotopic (i.e., hardware, operating system, Internet). Trained experts (such as graduate research assistants) also rated the 192 answers to the questions. The results of correlational analyses revealed that the LSA did an excellent job evaluating the quality of student answers. The correlation between LSA's answer quality scores and the mean quality scores of the experts was .49. This correlation is indistinguishable from the .51 correlation between the ratings of the two intermediate experts (i.e., the individuals who normally grade exams in a college computer literacy course). Graesser et al. (2000) reported that AutoTutor's LSA component did an excellent job discriminating the ability of learners who interact with AutoTutor in a multi-turn tutorial dialog. LSA was capable of discriminating different classes of student ability (good, vague, erroneous, versus mute students) and in tracking the quality of contributions in tutorial dialog.

The LSA space in AutoTutor was adopted in the present study. We computed the LSA-match scores between pairs of nodes in the conceptual graph structures that had been prepared for topics on hardware, operating systems, and the internet.

Conceptual Graph Structures on Topics in Computer Literacy

AutoTutor's architecture includes a set of conceptual graph structures on the various topics in the curriculum script. A typical structure contains approximately 10 to 30 nodes. We randomly selected 12 conceptual graph structures in the present analysis, including 4 structures for hardware, 4 for operating systems, and 4 for the internet.

The 12 knowledge structures were composed by applying the conceptual graph structure (CGS) representations developed by Graesser (Graesser & Clark, 1985; Graesser et al., 1992; Graesser, Wiemer-Hastings, & Wiemer-Hastings, in press; Williams, Hultman, & Graesser, 1998). The CGS's have 5 node categories: concepts, states, events, goals, and style specifications. There are 22 basic arc categories. The composition of these conceptual graph structures is not arbitrary, but is based on formal and conceptual constraints that have been studied for several decades in artificial intelligence (Lehmann, 1992). The categories of nodes and arcs are sufficient for implementing computational models of question answering which have been validated in experiments on adults (Baggett & Graesser, 1995; Graesser & Hemphill, 1991; Graesser, Lang, & Roberts, 1991).

Three types of knowledge structures were directly analyzed in the present study: taxonomic hierarchies, causal networks, and goal hierarchies. A node was included in the present analysis if and only if it was part of any of these three types of structures. The composition of these three types of structures is specified below.

Taxonomic Hierarchies

Concept nodes are connected by is-a arcs. For example, the concepts Norton Antivirus, utility program, and tool would be connected by two is-a arcs:

(concept-1: Norton Antivirus) –isa→
(concept-2: utility program) –isa→
(concept-3: tool)

The structure distance is 1 between concepts 1 and 2 and between concepts 2 and 3; the structural distance is 2 between concepts 1 and 3.

Causal Networks

State and event nodes are connected by arcs that signify Cause, Enables, Subprocess, and Implies (see Graesser & Clark, 1985 and Graesser, Wiemer-Hastings, & Wiemer-Hastings, in press for more complete definitions of arcs). Some of these categories of nodes and arcs are illustrated in the following chain.

(state-1: the operating system is stored on the hard disk) –Enable→
(event-2: the operating system is loaded onto the computer) –Subprocess→
(event-3: the operating system gets into RAM) –Cause→
(event-4: the CPU executes instructions)

The structural distance is 1 between nodes 1&2, 2&3, and 3&4, is 2 between nodes 1&3 and 2&4, and is 3 between nodes 1&4.

Goal-structures

Goal nodes are connected to other nodes by virtue of arcs that signify Reason, Manner, Initiate, and Outcome. For example, the following three goal nodes form a goal hierarchy via a Reason arc.

(goal-1: user types in command) –Reason→
(goal-2: user starts word processing software) –Reason→
(goal-3: user writes article)

The goals are triggered by various events and states in the world by virtue of Initiate arcs, whereas Outcome arcs specify whether or not the goals are achieved.

Scaling of Pairs of Nodes on Structural Proximity

Pairs of nodes in the 12 conceptual graph structures were scaled on structural proximity with respect taxonomic hierarchies, causal networks, and goal structures. A node in a structure was included in the analysis if and only if it was part of one or more of these three types of structures. When considering all 12 conceptual graph structures, there were 536 pairs of nodes in the analysis. A pair of nodes (A and B) was scaled on causal proximity by computing the reciprocal of the structural distance on a legal causal path between A and B (i.e., 1/distance). Thus, if two nodes have a structural distance of 1, 2, 3, versus 4 arcs on a legal path, then the causal proximity scores would be 1.00, .50, .33, and .25, respectively. If there is no legal causal path that connects A and B, the causal proximity score is 0. Goal proximity and taxonomic proximity was computed in a similar fashion for all 536 nodes. The mean proximity scores were .07, .31, and .40 for the taxonomic, causal, and goal proximity scores, respectively; the corresponding standard deviations were .26, .40, and .45.

Relationship Between LSA Match Scores and Structural Proximity Scores

The analyses uncovered a robust relationship between the LSA match scores and the structural proximity scores. Consider first the causal proximity scores. The mean LSA match scores were .47, .35, and .24 when the causal proximity scores were 1.00, .50, and .33 or lower (but not 0), respectively. When analyzing the goal proximity scores, the LSA match scores were .53 and .42 for goal proximity scores of 1.00 and .50 or lower (but not 0), respectively. The taxonomic proximity scores rarely went lower than 1.00 when considering nonzero values, so we could not isolate a sensitive gradient for this proximity score. The overall mean LSA match score was .44 (SD = .30).

A multiple regression was conducted to assess the extent to which the LSA match scores could be predicted by the taxonomic, causal, and goal proximity scores. The three predictor variables together explained a significant 9% of variance in the LSA match scores, $F(3, 532) = 16.46$, $p < .05$, $R^2 = .09$. All three predictors had a significant unique impact on the LSA scores, with beta weights of .14, .31, and .47 for taxonomic, causal, and goal proximity, respectively.

We performed some follow-up multiple regression analyses that statistically controlled for some potential extraneous variables. One extraneous variable was the length of the node descriptions, as defined by the number of words in the pair of nodes. Those who have conducted research on LSA have reported that lengthier descriptions have a slight tendency to produce higher LSA matches when two bags of words are compared (Rheder, Schreiner, Wolfe, Laham, Landauer, & Kintsch, 1998; Wiemer-Hastings et al., 1999). The mean length of the node descriptions in our sample was 10.62 words in the node pair (SD = 4.03), or 5.31 words per node. A second extraneous variable was the number of nouns that overlap between the pair of nodes. Overlapping nouns are analogous to argument overlap in propositional theories of text processing (Graesser, Millis, & Zwaan, 1997; Kintsch, 1998); the fact that constituents refer to the same entity is one important foundation for coherence in discourse processing. However, from the standpoint of the present analyses, we would not be particularly surprised if the LSA match scores could be explained by mere noun overlap because it is analogous to a keyword overlap. The mean number of overlapping nouns in a node pair was .71 (DS = .67).

Table 1 presents the results of the multiple regression analysis that predicted LSA match scores as a function of the three structural proximity scores, length, and noun overlap. The five predictor variables accounted for a significant 55% of the variance in LSA match scores, , $F(5, 530) = 128.05$, $p < .05$, $R^2 = .55$. When considering the two extraneous variables, noun overlap had a robust impact on the LSA match scores whereas length had no significant effect. Although noun overlap was robust, the three structural proximity variables still had a significant unique impact on the LSA match scores in the multiple regression analyses. Interestingly, we did not find the noun overlap scores to be correlated very highly with the taxonomic, causal, and goal proximity scores, r = -.18, .25, and -.02, respectively. Overlap in predicates was also analyzed but the correlations were also modest or nonsignificant. These results support the claim that the structural proximity scores have an impact on LSA match scores over and above keyword matches.

Table 1: Multiple regression analyses that predict LSA match scores

Predictor Variable	beta-weight	t-score
Taxonomic proximity	.14	4.08 *
Causal proximity	.11	2.17 *
Goal proximity	.15	2.94 *
Length (number of words)	-.02	.75
Noun overlap	.72	23.20 *

* significant at $p < .05$.

Conclusions

The results of this study support the claim that LSA captures aspects of the mental models that underlie computer literacy. The content of the mental models includes taxonomic structures, causal networks, and goal/plan/action hierarchies. The LSA match scores between pairs of nodes in the conceptual graph structures can be predicted by taxonomic, causal, and goal structural proximity. The structural proximity scores predict LSA match scores over and above noun overlap, keyword overlap, and the number of words in the node descriptions.

Aside from demonstrating that LSA captures aspects of mental models, we have demonstrated that LSA can be useful for performing semantic and conceptual analyses on relatively short verbal descriptions. Researchers have sometimes claimed that LSA is only useful when analyzing lengthier verbal descriptions on the order of a paragraph. The present study supports the claim that LSA can be useful for compositional analyses on individual words and short sentences of 5-6 words. Additional research is needed to identify the limits of LSA in recovering different aspects of semantics and world knowledge.

Acknowledgements

This research was funded by the National Science Foundation (SBR 9720314) in a grant awarded to the first author of this manuscript. The following members of the Tutoring Research Group at the University of Memphis conducted research on this project: Ashraf Anwar, Laura Bautista, Myles Bogner, Tim Brogdon, Patrick Chipman, Scotty Craig, Rachel DiPaolo, Evan Drumwright, Stan Franklin, Max Garzon, Barry Gholson, Art Graesser, Doug Hacker, Peggy Halde, Derek Harter, Jim Hoeffner, Xiangen Hu, Jeff Janover, Ashish Karnavat, Bianca Klettke, Roger Kreuz, Kristen Link, Shulan Lu, Johanna Marineau, William Marks, Lee McCaulley, Brent Olde, Para Orfanides, Natalie Person, Victoria Pomeroy, Penelope Price, Sonya Rajan, Akshay Thota, Mat Weeks, Holly Yetman White, Shannon Whitten, Katja Wiemer-Hastings, Peter Wiemer-Hastings, Shonijie Yang, and Zhaohua Zhang.

References

Baggett, W.B., & Graesser, A.C. (1995). Question answering in the context of illustrated expository text. *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 334-339). Hillsdale, NJ: Lawrence Erlbaum.

Burgess, C., Livesay, K., & Lund, K. (1998). Exploring context space: Words, sentences, discourse. *Discourse Processes*, *25*, 211-257.

Charniak, E. (1993). __*Statistical language analysis*. Cambridge, MA: Cambridge University Press.

Chi, M.T.H., Slotta, J.D., & de Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction, 4*, 27-43.

Foltz, P.W. (1996). Latent semantic analysis for text-based research. *Behavior Research Methods, Instruments, and Computers*, *28*, 197-202.

Foltz, PW., Britt, M.A., & Perfetti, C.A. (1996). Reasoning from multiple texts: An automatic analysis of readers' situation models. In G. Cottrell (Ed.), *Proceedings of the 18th Annual Cognitive Science Conference* (pp. 110-115). Mahwah, NJ: Erlbaum.

Forbus, K., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*, 141-205.

Golden, R.M. (1997). Causal network analysis validation using synthetic recall protocols. *Behavior Research Methods, Instruments, and Computers, 29*, 15-24.

Graesser, A.C., & Clark, L.C. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Ablex.

Graesser, A.C., Franklin, S., & Wiemer-Hastings, P. and the TRG (1998). Simulating smooth tutorial dialog with pedagogical value. *Proceedings of the American Association for Artificial Intelligence* (pp. 163-167). Menlo Park, CA: AAAI Press.

Graesser, A. C., Gordon, S. E., & Brainerd, L. E. (1992). QUEST: A model of question answering. *Computers and Mathematics with Applications, 23*, 733-745.

Graesser, A. C. & Hemphill, D. (1991). Question answering in the context of scientific mechanisms. *Journal of Memory and Language, 30*, 186-209.

Graesser, A. C., Lang, K. L., & Roberts, R. M. (1991). Question answering in the context of stories. *Journal of Experimental Psychology: General, 120*, 254-277.

Graesser, A.C., Millis, K.K., & Zwaan, R.A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163-189.

Graesser, A.C., Wiemer-Hastings, P., & Wiemer-Hastings, K. (in press). Constructing inferences and relations during text comprehension. In T.Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects*. Amsterdam: Benjamins.

Graesser, A.C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., and the TRG (in press). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research*.

Graesser, A.C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., and the TRG (2000). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. *Interactive Learning Environments*.

Kiel, F.C. (1979). *Semantic and ontological development: An ontological perspective*. Cambridge, MA: Harvard University Press.

Kintsch, E., Kintsch, W., Laham, D., Landauer, T.K., DePaula, R., Schreiner, M.E., Stahl, G., & Steinhart, D. (2000). Learning how to summarize using LSA-based feedback. *Interactive Learning Environments*.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, MA: Cambridge University Press.

Landauer, T.K., & Dumais, S.T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211-240.

Landauer, T.K., Foltz, P.W., Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*, 259-284.

Lehmann, F. (1992)(Eds.). *Semantic networks in artificial intelligence*. New York: Pergamon.

Lenat, D.B. (1995). CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM, 38*, 33-38.

Miller, G.A., & Johnson-Laird, P.N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.

Norman, D.A., & Rumelhart, D.E. (1975). *Explorations in cognition*. San Francisco, CA: Freeman.

Perfetti, C.A. (1998). The limits of cooccurrence: Tools and theories in language research. *Discourse Processes*, *25*, 363-377.

Rheder, B., Schreiner, M.E., Wolfe, M.B.W., Laham, D., Landauer, T.K., & Kintsch, W. (1998). Using latent semantic analysis to assess knowledge: Some technical considerations. *Discourse Processes, 25*, 337-354.

Schank, R.C., & Abelson, R. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.

Sowa, J.F. (1983). *Conceptual structures: Information processing in mind and machine*. Reading, MA: Addison-Wesley.

Trabasso, T., van den Broek, P.W. & Suh, S. (1989). Logical necessity and transitivity of causal relations in the representation of stories. *Discourse Processes, 12*, 1-25.

Wiemer-Hastings, P., Graesser, A.C., Harter, D., and the TRG (1998). The foundations and architecture of AutoTutor. *Proceedings of the 4th International Conference on Intelligent Tutoring Systems* (pp. 334-343). Berlin, Germany: Springer-Verlag.

Wiemer-Hastings, P., Wiemer-Hastings, K., and Graesser, A. (1999). Improving an intelligent tutor's comprehension of students with Latent Semantic Analysis. *Artificial Intelligence in Education* (pp. 535-542). Amsterdam: IOS Press.

Williams, K.E., Hultman, E., & Graesser, A.C. (1998). CAT: A tool for eliciting knowledge on how to perform procedures. *Behavior Research Methods, Instruments, & Computers, 30*,565-572.

Wolfe, M.B.W., Schreiner, M.E., Rehder, B., Laham, D., Foltz, P.W., Kintsch, W., Landauer, T.K. (1998). Learning from text: Matching readers and texts by latent semantic analysis. *Discourse Processes, 25*, 309-336.