Analyzing Human Tutorial Dialogues for Cohesion and Coherence During Hypermedia Learning of a Complex Science Topic

Moongee Jeon (MJEON1@memphis.edu)
Department of Psychology | Institute for Intelligent Systems | University of Memphis,
3693 Norriswood Avenue
Memphis, TN 38152 USA

Roger Azevedo (RAZEVEDO@memphis.edu)
Department of Psychology | Institute for Intelligent Systems | University of Memphis,
3693 Norriswood Avenue
Memphis, TN 38152 USA

Abstract

We analyzed cohesion and coherence of 79 think-aloud transcription logs from a human tutorial dialogue study investigating the effect of tutoring on college students' learning about the circulatory system with hypermedia. The study involves randomly assigning 82 non-science majors to either the self-regulated learning (SRL, control condition) or the externally-regulated learning (ERL, experimental condition with a human tutor). The corpus we examined contained a total of 1,445 pages. We used Coh-Metrix, an automated web-based tool developed to evaluate text and discourse, in order to assess cohesion and coherence of the tutorial dialogues produced between a human tutor and low-domain knowledge college students during 79 tutoring sessions. Our findings showed that there were significant differences in the tutorial dialogues of the ERL tutoring condition versus those of the SRL control condition in the co-referential cohesion, the semantic/conceptual overlap, the causal ratio, the standard readability formulas, the incidence scores of connectives, the number of words, the number of sentences, the number of turns, the average sentences per turn, and the average words per sentence. Our findings have implications for the design of intelligent tutorial dialogue in hypermedia systems developed to improve learners' deep conceptual understanding of complex and challenging science topics.

Keywords: Cohesion; Coherence; Human Tutorial Dialogue; Learning; Hypermedia; Human Tutoring

Cohesion and Coherence in Tutorial Dialogues

Cohesion and coherence of text and discourse are critical components affecting text processing and comprehension. According to Graesser et al. (2004), coherence indicates characteristics of mental models that readers establish during comprehension, whereas cohesion refers to features of text-based information such as argument overlap, discourse markers, anaphora, and connectives. Coherence, according to Graesser et al. (2003), is the final psychological construct that readers build in their mind. Readers normally construct a coherent representation in terms of various cohesion and coherence relations during comprehension. For example, they attain coherence of a text by identifying various coherence relations (Sanders, Spooren, & Noordman, 1992) or rhetorical structures (Mann & Thompson, 1988) in the text, linking text-based segments, and combining the text-based information with their prior background knowledge (Kintsch, 1988, 1998; McNamara et al., 1996). Similarly, students can construct coherent mental models for subject matters while learning with the help of a human tutor in a classroom or while interacting with a computer-based learning environment such as a hypermedia environment (e.g., Azevedo et al., 2004, 2005, in press).

Recent empirical studies have shown that cohesion and coherence are cardinal to examine whether readers generate inferences to link text units during on-line comprehension (Graesser, Singer, & Trabasso, 1994; Graesser & Bertus, 1998; Millis & Graesser, 1994), to combine the text-based information with the readers' prior knowledge in their long-term memory (McNamara et al, 1996), and to construct mental models successfully in their mind (Gernsbacher, 1997; Graesser, Singer, & Trabasso, 1994; Kintsch, 1988, 1998; Zwaan, Langston, & Graesser, 1995). The mental models established by the readers promote deeper processing and understanding of the text. Thus, the coherent mental representation constructed based on the text-based information together with the readers' prior knowledge is essential to understand the text successfully at a deep level. While these indices have been adopted mainly by cognitive psychologists in text comprehension studies, they are now being used by members of the Cognitive Science community to examine the quality and linguistic features found in tutoring dialogues. The aim is to understand the nature of the student-tutor dialogues in order to better design automated conversational dialogue in learning technology-based systems aimed at fostering conceptual understanding of complex and challenging science topics (e.g., see Azevedo, 2005; Graesser, McNamara, & VanLehn, 2005).

With these theoretical frameworks and recent advance in computational linguistics and natural language technologies, our research group has developed Coh-Metrix, an automated web-based tool to analyze cohesion and coherence of various types of text, discourse, and tutorial dialogue in terms of over 400 measures (Graesser, McNamara, Louwerse, & Cai, 2004). The Coh-Metrix tool has been widely used by our research group. For example, the Coh-Metrix tool was recently used by Jeon and Graesser (2006) to examine cohesion, coherence, and readability of the tutorial dialogues of high-knowledge students who had taken college physics versus the tutorial dialogues of low-knowledge students with no college physics background while interacting with AutoTutor, an animated pedagogical agent which scaffolds college students to learn about conceptual physics by holding conversations in natural language (Graesser, Person, Lu, Jeon, & McDaniel, 2005). Their results showed that the tutorial dialogues of the high-knowledge students were connected more coherently than of the low-knowledge students.

Coh-Metrix has been used, in this human tutoring study, to examine cohesion, coherence, and readability of the tutorial dialogues of the ERL (externally-regulated learning) tutoring condition versus the tutorial dialogues of the SRL (self-regulated learning) control condition.

Coh-Metrix

Coh-Metrix (http://cohmetrix.memphis.edu/) is an automated web-based tool designed to examine text and discourse on cohesion, coherence, and readability based on over 400 measures (Graesser, McNamara, Louwerse, & Cai, 2004). For example, the Coh-Metrix tool provides co-referential cohesion (argument overlap and stem overlap), LSA cosines (latent semantic analysis: semantic/conceptual overlap), causal ratio, incidence scores of connectives and logical operators, and other measures. Coh-Metrix examines texts using various types of cohesion and language characteristics unlike standard readability formulas (e.g., Flesch Reading Ease and Flesch-Kincaid Grade Level) which mainly focus on word length and sentence length. Coh-Metrix consists of lexicons, part-of-speech classifiers, syntactic parsers, semantic templates, latent semantic analysis, and other components that have been extensively used in computational linguistics.

Method: An Overview of the Study

The overall study conducted by Azevedo and colleagues (in press) examined the effectiveness of self-regulated learning (SRL) and externally regulated learning (ERL) on college students' learning about a science topic with hypermedia during a 40-minute session. A total of 82 college students with little knowledge of the topic were randomly assigned either to the SRL or ERL condition. Students in the SRL condition regulated their own learning, while students in the ERL condition had access to a human tutor who facilitated their self-regulated learning. We converged product (pretest-posttest declarative knowledge and qualitative shifts in participants' mental models) with process (think-aloud) data to examine the effectiveness of SRL versus ERL. Analysis of the declarative knowledge measures showed that the ERL condition group mean was statistically significantly higher than the group mean for the SRL condition on the labeling and flow diagram tasks. There were no statistically significant differences between groups on the matching task, but both groups showed statistically significant increases in performance. Further analyses showed that the odds of being in a higher mental model posttest group were decreased by 65% for the SRL group as compared to the ERL group. In terms of self-regulatory processes, participants in the SRL condition engaged in more use of less effective learning strategies (e.g., free searching), while those in the ERL participants engaged in more planning and activation of prior knowledge, deployment of key monitoring activities (e.g., monitoring their progress toward goals), and effective strategies (e.g., hypothesizing).

The goal of this paper is to report the analyses that were conducted on the 79[1] think-aloud protocols (41 from the SRL control condition and 38 from the ERL tutoring condition) by focusing on the cohesion and coherence measures extracted from the 1,445 pages of think-aloud protocols.

The participants were non-science majors from a large mid-Atlantic public university who were given extra credit to participate in the study. Their mean age was 21 with an average GPA (mean = 3.3) and comprised mostly of female students (81% of the sample).

Tutoring Sessions

Briefly, the methodology of the original study which has been extensively used by Azevedo and colleagues (see Azevedo, Cromley, Winters, Moos, & Greene, 2005; Azevedo, Moos, Greene, Winters, & Cromley, in press; Greene, Moos, Azevedo, & Winters, 2006; Moos & Azevedo, 2006) was also used in this study to evaluate students' self-regulated learning about science with hypermedia: (1) obtain informed consent; (2) administer participant questionnaire; (3) administer pretest (20 minutes); (4) provide instructions for the learning task; (5) tour of the hypermedia environment including its features and content; (6) give participants think-aloud practice task; (7) inform participants of the overall learning goal ("*Make sure you learn about the different parts and their purpose, how they work both individually and together, and how they support the human body*") as part of their instructions for learning about the circulatory system. During the 40-minute hypermedia learning task, the participants had access to the think-aloud instructions and the overall learning goal. Participants were free to create drawings or take notes while learning with hypermedia, although not all chose to do so. All participants were required to navigate the hypermedia environment for 40 minutes to learn about the circulatory system. Differences between the two conditions (ERL versus SRL) were given next. Only the participants in ERL had access to a human tutor with extensive training (B.Sc. in biology and six years of teaching biology in the schools) who would scaffold them to learn about the circulatory system by furnishing external regulation and by prompting them throughout the session to activate their prior knowledge, to establish learning goals, to monitor several aspects of their learning and task conditions, to deploy key learning strategies (e.g., summarize, coordinate informational sources), and to handle task difficulties and demands (e.g., engage in help-seeking behavior).

[1] We dropped 3 participants because of poor audio quality.

Materials and Procedures for Analyzing Tutorial Dialogue Data

We collected 38 ERL transcription logs and 41 SRL transcription logs from the 79 college students who participated in this human tutoring study all of which were the Microsoft Office Word files. Here are examples of the ERL and SRL transcriptions:

ERL Transcription (Subject #3030)
Tutor: Scroll down till you find the red blood cell section, just read the first sentence. You'll find out.
Student: Carry oxygen from the lungs.
> READS: Red blood cells make up almost 45 percent of the blood volume…

Tutor: Yeah. They have a very special molecule called hemoglobin and it carries oxygen. They are the, they are the key to the oxygen.
Student: lungs, lungs…

SRL Transcription (Subject #3063)
Student: yeah …
> …
> READS: This inner layer is surrounded by connective tissue and smooth muscle. Blood vessels expand during exercise...

… this picture here... constituents of blood, an average healthy person approximately 45% of blood volume in the cells... clear yellow layer called plasma so plasma's yellow. hmm... so... I guess the red part would just be the red blood cells... ummm... 95% of plasma's water and glucose, proteins, amino acids... hmm… [unintelligible]…

To analyze cohesion and coherence in the human tutorial dialogues with the Coh-Metrix tool, we extracted only the tutor and student turns from the 79 think-aloud transcription logs, excluding AUDIO VIDEO and READS parts of the transcriptions. We then eliminated all the titles such as 'Student' and 'Tutor' and the bracketed transcriptions (e.g., [unintelligible]) in the tutorial dialogues. Next, we divided the whole tutorial dialogues including both the tutor and student turns into two types of dialogue group (tutor versus student dialogue group) because the primary research question of this paper is to investigate cohesion, coherence, and readability of the tutorial dialogues of the ERL tutoring condition versus the SRL control condition using Coh-Metrix. We used Coh-Metrix (version 1.4) entering the three types of tutorial dialogue: the whole tutorial dialogue, the tutor dialogue, and the student dialogue. However, we analyzed only the students' dialogues for this paper.

Results

For this paper, we investigated a few specific measures of the Coh-Metrix tool. More specifically, we selected the measures such as the co-referential cohesion scores, the latent semantic analysis, the standard readability formulas, the causal ratio, the incidence score of connectives, the number of words, the number of sentences, the number of turns, the average words per sentence, and the average sentences per turn because the main goal of this paper was to compare cohesion, coherence, and readability of the tutorial dialogues from the ERL (externally-regulated learning) condition with those from the SRL (self-regulated learning) control condition while learning with hypermedia. The results presented in this section are from several one-way ANOVAs performed between the means of several Coh-Metrix measures of the ERL and the SRL conditions.

Co-Referential Cohesion Scores

Co-reference cohesion occurs when a noun, pronoun, or noun phrase refers to another constituent in discourse. For this paper, the Coh-Metrix tool computed two types of co-reference cohesion scores including the adjacent argument overlap and the stem overlap scores. The adjacent argument overlap score, in this paper, is defined as the proportion of adjacent sentence pairs in a tutorial dialogue that share one or more arguments and the adjacent stem overlap indicates the proportion of adjacent sentences that share one or more stems (e.g., **swim**, **swim**mer, **swim**ming).

Table 1: Means (*SDs*) for the measures of Coh-Metrix by Tutoring Condition

Measures of Coh-Metrix	ERL Condition (n=38)	SRL Condition (n=41)
Adjacent Argument Overlap	.15(.09)	.36(.14)
Adjacent Stem Overlap	.12(.08)	.33(.13)
LSA Cosine of Adjacent Sentence to Sentence	.18(.05)	.24(.07)
LSA Cosine of Turn to Turn	.15(.1)	.41(.2)
Flesch-Kincaid Grade Level (0-12)	3.1(1.3)	7.5(2.7)
Flesch Reading Ease Score (0-100)	83(6.8)	66(11.4)
Causal Ratio	3.4(1.7)	2.1(1.4)
Incidence Score of Total Connectives	89(18.8)	74(21.8)
Incidence Score of Positive Causal Connective	30.9(9)	22.9(10.6)
Incidence Score of Negative Additive Connective	5.8(3.1)	2.8(2.6)
Number of Words	1323(457)	1083(594)
Number of sentences	241(84)	92(62)
Number of Turns	177(54)	13(9)
Average Sentences per Turn	1.4(.2)	10.1(7.8)
Average Words per Sentence	5.8(2.2)	13.4(6.2)

Two separate ANOVAs were performed on each of the adjacent argument overlap and the adjacent stem overlap scores of tutorial dialogues as a function of condition (ERL verses SRL; see Table 1). Our findings showed that there was a significant difference between the ERL tutoring condition and the SRL control condition for the adjacent argument overlap, F (1, 77) = 64.7, MSE = .013, p < .001; the argument overlap score of the SRL condition (M = .36) was higher than the ERL condition (*M* = .15). Similarly, the stem overlap score of the SRL tutorial dialogues (*M* = .33) was also significantly higher than that of the ERL (*M* = .12) [F (1, 77) = 71.8, *MSE* = .012, $p < .001$] (see Table 1). These results show that the tutorial dialogues of the SRL control condition are more co-referential than the ERL tutoring condition indicating that the adjacent sentences of the SRL condition dialogues have more common arguments (e.g., noun, pronoun, noun-phrase) and stems than of the ERL tutoring condition.

LSA Cosine of Adjacent Sentence to Sentence

Latent Semantic Analysis (LSA) is a mathematical and statistical method which computes semantic similarities of sentences and paragraphs of text and discourse (Landauer, Foltz, & Laham, 1998). An LSA cosine of adjacent sentence to sentence, in this paper, indicates semantic relatedness of adjacent sentences in tutorial dialogues. In general, high LSA cosines of adjacent sentence to sentence represent that the sentences are connected more semantically. Here is an example of a student's tutorial dialogue taken from the SRL tutoring condition (Subject #3063) to show how the LSA cosines of adjacent sentences were computed.

> Turn: Four chambers, left right ventricle, left right atrium (Sentence1). The muscle is the myocardium (Sentence2)…
> Turn: … lets see we have the three types of blood vessels, arteries, veins, capillaries (Sentence 3). Arteries away veins toward I remember that much (Sentence 4) …

The mean of LSA cosines for adjacent sentence to sentence was computed as follows:

> LSA cosine of adjacent sentence to sentence = MEAN [LSA (Sentence 1, Sentence 2), LSA (Sentence 2, Sentence 3), LSA (Sentence 3, Sentence 4)]

An ANOVA was performed on the LSA cosines of adjacent sentence to sentence as a function of condition (ERL versus SRL; see Table 1). There was a significant difference between the ERL tutoring condition and the SRL control condition, F (1, 77) = 39.8, *MSE* = .005, $p < .001$. The LSA cosine of adjacent sentence to sentence of the SRL tutorial dialogues (*M* = .24) was higher than of the ERL tutoring condition (M = .18) indicating that the coherence of the adjacent sentences of the SRL tutorial dialogues is higher than of the ERL in the local level (means local coherence).

LSA Cosine of Turn to Turn

LSA cosines of turn to turn indicate semantic relatedness of turns in tutorial dialogues. For example, high LSA cosines of turn to turn indicate that the turns in the tutorial dialogues are linked more semantically. Here is an example of students' tutorial dialogues taken from the SRL tutoring condition (Subject #3063) to show how the LSA cosine of turn to turn was calculated.

> Turn1: The heart... superior, okay so there's the vena cava, superior and inferior, they go into the right atrium, tricuspid valve, right ventricle. uhh pulmonary...
> Turn2: ... I was using them actually right. so umm... proteins, protein molecules, lymphocytes... the immune system, fraction of the blood composition, gamma, ...
> Turn3: ... umm blood, vital fluid, river of life, that's not really useful. Role of blood, lungs, digestive system, …

The mean of LSA cosines of turn to turn was calculated as follows:

> LSA cosine of turn to turn = MEAN [LSA (Turn 1, Turn 2), LSA (Turn 1, Turn 3), LSA (Turn 2, Turn 3)]

We performed an ANOVA on the LSA cosines of turn to turn as a function of condition (see Table 1). Our results showed that the LSA cosines of turn to turn in the SRL condition (*M* = .41) were significantly higher than those in the ERL tutoring condition (*M* = .15), F (1, 77) = 58.3, *MSE* = .022, $p < .001$. This result indicates that the turns of the tutorial dialogue of SRL control condition are connected more coherently than of the ERL tutoring condition in the global level (means global coherence).

Standard Readability Formulas

The Coh-Metrix tool provides two types of standard readability formulas, the Flesch Reading Ease score and the Flesch-Kincaid Grade Level. The Flesch Reading Ease score is expressed by a number ranging from 0 to 100; a higher score represents easier reading. The Flesch Reading Ease score is computed as follows:

> Flesch Reading Ease Score = 206.835 - (1.015 * ASL) - (84.6 * ASW)

> ASL (Average Sentence Length) = the number of words divided by the number of sentences.
> ASW (Average number of Syllables per Word) = the number of syllables divided by the number of words.

The Flesch-Kincaid Grade Level defined as a number ranging 0 to 12 converts the Flesch Reading Ease Score to a U.S. grade-school level; a higher score indicates more difficult reading. The Flesch-Kincaid Grade Level is computed as follows:

> Flesch-Kincaid Grade Level = (.39 * ASL) + (11.8 * ASW) - 15.59

We performed two separate ANOVAs on the two standard readability formulas as a function of condition (ERL versus SRL). The results showed that there was a significant difference between the ERL tutoring condition and the SRL control condition for the Flesch Reading Ease score, F (1, 77) = 68.1, MSE = 89.8, p < .001. The Flesch Reading Ease score of the ERL condition (M = 83) was higher than the SRL condition (M = 66). There was also a significant difference between the two conditions for the Flesch-Kincaid Grade Level, F (1, 77) = 81.8, MSE = 4.5, p < .001. The Flesch-Kincaid Grade Level of the SRL condition (M = 7.5) was higher than of the ERL condition (M = 3.1). These results suggest that the sentences and turns of the tutorial dialogues of the SRL condition exhibit more complicated syntactic structures than of the ERL condition indicating more difficult reading (see Table 1).

Causal Ratio

Causal cohesion indicates the causal relatedness of sentences in terms of causal cohesion relations (Graesser et al., 2004). The Coh-Metrix tool calculates causal cohesion relations by causal ratios. A causal ratio is defined as a proportion of causal particles (P) to causal verbs (V). Simply put, the causal ratio is computed as follows:

Causal Ratio = P / (V+1)

The denominator was the number of causal verbs plus one in order to consider the rare case in which a text has no causal verbs. An ANOVA was performed on the causal ratios of the tutorial dialogues as a function of condition. The result showed that the causal ratio of the ERL condition (M = 3.4) was significantly higher than the SRL control condition (M = 2.1), F (1, 77) = 13.9, MSE = 2.4, p < .001, reflecting that the sentences and turns of the tutorial dialogues of the ERL tutoring condition are connected more causally based on the causal relations signaled by the causal particles (e.g., *since*, *so that*, and *consequently*) and verbs (e.g., *cause*, *enable*, and *make*) than of the SRL control condition (see Table 1).

Incidence Score of Connectives

The Coh-Metrix tool computes incidence scores of several types of connectives in discourse such as temporal, additive, and causal connectives. An incidence score is expressed by a number of occurrences of a particular category per 1,000 words. ANOVAs were performed on the incidence scores of connectives in the tutorial dialogues as a function of condition. Our results showed that the incidence score of the total connectives of the ERL condition (M = 89) was significantly higher than the SRL control condition (M = 74), F (1, 77) = 11.2, MSE = 416.7, p = .001. More specifically, students in the ERL tutoring condition (M = 30.9) used more positive causal connectives (e.g., *therefore* and *thus*) than in the SRL control condition (M = 22.9), F (1, 77) = 13.2, MSE = 96.5, p < .001. The students in the ERL tutoring condition (M = 5.8) also used more negative additive connectives (e.g., *however* and *but*) than in the SRL control condition (M = 2.8), F (1, 77) = 21.3, MSE = 8.2, p < .001. These results indicate that the students in the ERL tutoring condition tended to use more positive causal and negative additive connectives in order to gain a deep conceptual understanding (based on mental model shifts of the circulatory system from pretest to posttest) while interacting with the human tutor (see Table 1).

Basic Count

The Coh-Metrix tool computed the number of words, the number of sentences, the number of turns, the average sentences per turn, and the average words per sentence in the tutorial dialogues. ANOVAs were performed on each of them as a function of condition. Our results showed that (1) the number of words in the tutorial dialogues of the ERL condition (M = 1,323) was significantly higher than those in the SRL condition (M = 1083), F (1, 77) = 4.02, MSE = 283171.4, p < .05; (2) the number of sentences in the ERL transcriptions (M = 241) was significantly higher than those in the SRL condition (M = 92), F (1, 77) = 81.3, MSE = 5368.9, p < .001; and (3) the number of turns in the ERL condition (M = 177) was significantly higher than those in the SRL condition (M = 13), F (1, 77) = 372, MSE = 1426.8, p < .001. By contrast, the average sentences per turn of those in the SRL condition (M = 10.1) was significantly higher than those in the ERL condition (M = 1.4), F (1, 77) = 48, MSE = 31.8, p < .001, and the average words per sentences in the SRL condition (M = 13.4) was significantly higher than those in the ERL condition (M = 5.8), F (1, 77) = 52.2, MSE = 21.1, p < .001. These results indicate that the students in the ERL tutoring condition interacted more with a tutor than in the SRL control condition; which is normal given that the students in the SRL condition did not have access to a human tutor while learning about the circulatory system with hypermedia. By contrast, the SRL students created longer sentences, but with fewer turns than the ERL students. The longer sentences for SRL students might be compensating for the fewer number of turns during the learning session, but these linguistic features were not reflected in learning.

Summary of Results

In sum, we analyzed cohesion and coherence from 79 think-aloud transcription logs (1,445 pages) from a human tutoring study investigating the effect of tutoring on college students' learning about the challenging and complex science topic with hypermedia. We used Coh-Metrix to examine cohesion and coherence of the tutorial dialogues produced between a human tutor and low-domain knowledge college students during 79 tutoring sessions. Our findings showed that there were significant differences in the tutorial dialogues of the ERL (externally-regulated learning) tutoring condition versus the SRL (self-regulated learning) control condition in the co-referential cohesion, the semantic/conceptual overlap, the causal ratio, the standard readability formulas, the incidence scores of connectives, the number of words, the number of sentences, the number of turns, the average sentences per turn, and the average words per sentence. Considering the novelty of these methods within the Cognitive Science community, we argue that our findings have implications for the design of intelligent tutorial dialogue in hypermedia systems developed to promote learners' deep conceptual understanding of complex and challenging science topics.

Acknowledgments

This research was supported by funding from the National Science Foundation (REC#0133346 and REESE#0633918) awarded to the second author. The authors would like thank Jeffrey Greene, Daniel Moos, Dr. Cromley, Fielding Winters, and Neil Hofman for assistance with data collection and transcribing the audio data. We also acknowledge Dr. Graesser's Institute for Education Sciences (R3056020018-02) and National Science Foundation (REC 0106965 and ITR 0325428) grants supporting the development of Coh-Metrix. We also thank Amy Witherspoon for comments and feedbacks on a previous draft.

References

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist, 40,* 199-209.

Azevedo, R., Cromley, J.G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29*, 344-370.

Azevedo, R., Cromley, J.G., Winters, F.I., Moos, D.C., & Greene, J.A. (2005). Adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instructional Science, 33,* 381-412.

Azevedo, R., Moos, D.C., Greene, J.A., Winters, F.I., & Cromley, J.C. (in press). Why is externally-regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development.*

Gernsbacher, M. A. (1997). Two Decades of Structure Building. *Discourse Processes, 23,* 265-304.

Graesser, A.C., & Bertus, E.L. (1998). The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading, 2,* 247-269.

Graesser, A.C., Louwerse, M.M., McNamara, D., Olney, A., Cai, Z., & Mitchell, H. (in press). Inference generation and cohesion in the construction of situation models: Some connections with computational linguistics. In F. Schmalhofer and C. Perfetti (Eds.), *Higher level language processes in the brain: Inferences and comprehension processes.* Mahwah, NJ: Erlbaum.

Graesser, A.C., McNamara, D.S., & Louwerse, M.M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text. In A.P. Sweet & C.E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82-98). NY: Guilford.

Graesser, A.C., McNamara, D.S., Louwerse, M.M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers, 36*, 193-202.

Graesser, A.C., McNamara, D., & VanLehn, K. (2005). Scaffolding Deep Comprehension Strategies through Point&Query, AutoTutor and iSTART. *Educational Psychologist, 40,* 225-234.

Graesser, A.C., Person, N., Lu, Z., Jeon, M., & McDaniel, B. (2005). Learning while holding a conversation with a computer. In L. PytlikZillig, M. Bodvarsson, & R. Bruning (Eds.), *Technology-based education: Bringing researchers and practitioners together* (pp. 143-167). Greenwich, CT: Information Age Publishing.

Graesser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371-95.

Greene, J.A., Moos, D.C., Azevedo, R., & Winters, F.I. (2006). Exploring differences between gifted and grade-level students' use of self-regulatory learning processes with hypermedia. *Computers & Education*.

Jeon, M., & Graesser, A. C. (November, 2006). Using Coh-Metrix to Assess the Role of Background Knowledge on Cohesion in Tutorial Dialogue. Symposium presentation at the 36th Annual Meeting of the Society for Computers in Psychology. Houston, TX.

Kintsch, W. (1988). The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model. *Psychological Review, 95,* 163-182.

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

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes, 25,* 259-284.

Mann, W.C., & Thompson, S.A. (1988). Rhetorical structure theory: towards a functional theory of text organization. *Text, 8*, 243-281.

McNamara, D.S., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14,* 1-43.

Millis, K. & Graesser, A.C. (1994). The time-course of constructing knowledge-based inferences for scientific texts. *Journal of Memory and Language, 33,* 583-599.

Moos, D.C., & Azevedo, R. (2006). The role of goal structure in undergraduates' use of self-regulatory variables in two hypermedia learning tasks. *Journal of Educational Multimedia and Hypermedia, 15(1),* 49-86.

Sanders, T.J.M., Spooren, W., & Noordman, L.G.M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes, 15,* 1, 1-35.

Zwaan, R.A., Langston, M.C., & Graesser, A.C. (1995). The construction of situation models in narrative comprehension: an event-indexing model. *Psychological Science, 6,* 292-297.

Zwaan, R.A., & Radvansky, G.A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162-185.