Why are computational models of text comprehension useful

Article in Revista Signos: Estudios de Lengua y Literatura · January 2007

3 authors:

Susan R. Goldman
University of Illinois at Chicago
208 PUBLICATIONS 7,610 CITATIONS

Richard M Golden
University of Texas at Dallas
103 PUBLICATIONS 1,698 CITATIONS

Paul van den Broek
Leiden University
196 PUBLICATIONS 13,024 CITATIONS

Goldman, S. R., Golden, R. M., & Van den Broek, P. (in press). Why are computational models of text comprehension useful? In F. Schmalhoefer & C. A. Perfetti (Eds.), Higher Level Language Processes in the Brain: Inference and Comprehension Processes. Mahwah, NJ: Erlbaum.

Why are Computational Models of Text Comprehension Useful?

Susan R. Goldman,[1] Richard M. Golden[2], & Paul van den Broek[3]

Text comprehension is a complicated process. Phenomena such as word perception, syntactical analysis, semantic analysis, and inference making are essential components of the text comprehension process. Not surprisingly, most empirical research and theories encompass only a subset of the phenomena and processes that constitute a complete account of text comprehension. Indeed, the component phenomena are themselves quite complicated and there are multiple competing theoretical accounts of them. Theoretical accounts of text comprehension are further complicated by the need to consider production of text. This is so because a large body of research assesses text comprehension via text that the comprehender produces, usually from memory. In the face of such complexity, many theories of text comprehension focus on a subset of the phenomena and attempt to create psychological process models that can account for behavioral data.

[1] Department of Psychology (MC 285), 1007 West Harrison Street, University of Illinois at Chicago, 1009 Behavioral Sciences Building, Chicago, Illinois 60607-7137. sgoldman@uic.edu

[2] School of Behavioral and Brain Sciences (GR4.1), Box 830688, University of Texas at Dallas, Richardson, TX 75083-0688. golden@utdallas.edu

[3] Educational Psychology, Burton Hall, 178 Pillsbury Drive, S. E. University of Minnesota, Minneapolis, MN55455. pvdbroek@tc.umn.edu

For the purposes of this book chapter, the term *model* refers specifically to a representation of the psychological processes that comprise a component or set of components involved in human text comprehension. *Computational models* refer to representations that are expressed in forms that can be run," providing simulated data that can be compared to data obtained from real people. Often computational models contain learning algorithms (e.g., the back propagation rule) and mathematical formalisms (e.g., global memory-matching – see Gillund & Shiffrin, 1984). that have been found to provide reasonably robust accounts of other learning and memory phenomena. Models of text comprehension, more so than models of simpler psychological phenomena, have benefited from the use of computers to run simulations because of the sheer computational power needed to capture the psychological complexity of text comprehension.

It is important to emphasize that a model is not equivalent to a theory. A *theory* is typically more comprehensive than any specific model and consists of a set of explicit assumptions about mechanisms and parameters, and logical arguments about the relations among them. Theories, especially of text comprehension, often permit the derivation of multiple models that differ in terms of the specific mechanisms and parameter values that they represent. Thus, multiple models might represent acceptable instantiations of the same theory. The process of generating multiple models and testing them is critical to the process of advancing theoretical accounts of text comprehension.

In this chapter we argue that computational models in particular have played an important role in the process of unraveling and understanding the psychological complexity of text comprehension. They have done so for three major reasons. First, the process of transforming verbally described theories of text comprehension (conceptual theories) into computational models of text comprehension promotes the development and evolution of the conceptual theories by showing where the models accord with behavioral data and where they do not. Agreements with behavioral data are evidence supporting the assumptions giving rise to the model whereas disagreements point out areas where the computational model, the theory or both need further development. Tests of alternative computational models further expand the usefulness of the enterprise for theory development. Second, computational models can be applied to behavioral data to better understand and test alternative explanatory constructs, especially in cases where patterns of behavioral data are not as expected *a priori.* In such cases, researchers provide post-hoc explanations, many of which are quite reasonable. Computational models can provide a way to test or enact such explanations. Because computational models make specific, and sometimes, non-obvious predictions, we can test alternative models against one another and the results can help us distinguish among competing conceptual theories. Finally, and partly as a result of the first two benefits, computational models promote communication among researchers within and across research areas. They promote consolidation and integration of theories and empirical findings about text comprehension, highlight areas where further theoretical development is needed, and integrate with other areas of research by showing where mechanisms important to text comprehension may also be important in understanding other phenomena. These claims are further developed and illustrated in the remainder of this chapter.

Computational Models Stimulate Theory Development

As discussed earlier, describing the psychological processes involved in text comprehension is complicated because a large number of cognitive systems are involved. Current theories of text comprehension acknowledge this complexity in assumptions about complex interactions among various levels and systems of language (e.g., words, sentences; syntax, semantics), especially in the face of limited attentional and verbal memory capacity resources (e.g., Gernsbacher, 1990; Goldman &Varma, 1995; Graesser, Singer, & Trabasso, 1994; Just & Carpenter, 1992; Kintsch, 1998; McKoon & Ratcliff, 1992; Myers & O'Brien, 1998; van Dijk & Kintsch, 1983; van den Broek, 1990). Advances in text comprehension theory have come about through efforts to translate theoretical formulations that posited such variables into tractable computational models. In showing what could and could not be accounted for, the computational efforts have spurred the evolution of text comprehension theories. We develop this position by first describing on the work of Walter Kintsch and colleagues because they have developed a very influential text comprehension theory in which computational modeling has played a major role. We then outline two other computational approaches that evolved from Kintsch's work.

Evolving Theories of Text Comprehension

The roots of a major class of current text comprehension theories can be traced to two seminal publications by Walter Kintsch, *The Representation of Meaning in Memory* (Kintsch, 1974) and a *Psychological Review* paper published in 1978 (Kintsch & van Dijk, 1978). In the former, Kintsch laid the groundwork for psychological theories of text processing and memory by documenting the linguistic and empirical research motivating the assumption that the proposition rather than the word or sentence is the appropriate unit for representing meaning. He showed systematic relations between propositions, reading time, and memory, using mathematical models of memory processes to account for data obtained when people read and recalled or recognized information from texts constructed to have specific characteristics. For example, three or four sentence paragraphs were written to contain the same number of words but different numbers of propositions. Propositional characteristics, such as the number of propositions in a sentence, were shown to predict behavioral data to a greater degree than did word characteristics, such as the number of words in a sentence (Kintsch, 1974).

In the 1978 paper, Kintsch and van Dijk (1978) proposed a theory of text processing that worked with propositional representations of the input text. In doing so, they consciously put aside issues of how people processed sentences to derive propositions. Rather they focused their theory on how propositions from successive sentences in a text were processed to produce connected and hierarchically organized sets of propositions. The 1978 theory was foundational for interactive models of comprehension and learning from text because it laid out a clear representational format for the text input; a processing model; and mechanisms for incorporating prior knowledge, comprehension goals, and strategies. It “located” comprehension in the interaction of the text, the reader, and the task, although at that time attention was primarily focused on the text.

Interactive theories of text comprehension (e.g., Gernsbacher, 1990; Goldman & Varma, 1995; Just & Carpenter, 1992; Kintsch, 1998; Myers & O’Brien, 1998; van Dijk & Kintsch, 1983; van den Broek, 1990) continue to dominate other classes of text comprehension models (e.g., letter by letter or word by word models such as that proposed by Gough (1972)). An important commonality among interactive models is that the on-line text comprehension process is assumed to proceed in a series of sequential cycles in which the reader processes a small group of propositions in each cycle, making connections among the new input and propositions from previous cycles. Details of the operation of cyclical processing differ somewhat among models. We use Kintsch and van Dijk's work (Kintsch & van Dijk, 1978) to illustrate the prototype and discuss how it evolved in response to results from both behavioral and computational modeling studies.

In the 1978 Kintsch and van Dijk model, the number of propositions processed on each cycle is a parameter, assumed to be equivalent to the number of chunks that can be held in working memory, typically 7 plus or minus 2. Note that the contents of a chunk are flexible and often vary across researchers and content domains. Connection making is subject to constraints imposed by a limited capacity verbal working memory so that not all previously processed propositions are typically available to connect with the new input. The number of propositions that is available from prior processing cycles when the next input is processed is termed the buffer size, represented as a parameter ***s.*** When new input fails to connect to available prior input, the reader reactivates previously processed propositions and/or makes connecting inferences based on prior knowledge. The results of the processing create an explicit text base representation (the set of propositions that were in the input) and an implicit textbase representation (the explicit textbase plus the propositions added through inference making during processing). The Kintsch and van Dijk (1978) theory posited that people often substitute a single proposition for several propositions, called a macroproprosition, but it was not until 1983 that the assumptions about the rules for generating macropropositions were laid out (van Dijk & Kintsch, 1983).

In the 1978 and 1983 versions of the theory, connection making is a critical process in achieving comprehension because it allows the propositional representation to reflect the semantic coherence across the sentences in the text. Coherence across sequential sentences is precisely what differentiates the processing of text as connected discourse from processing of lists of sentences. Although the 1978 theory discussed the importance of readers' goals and the task in relation to efforts to create coherence, little attention was given to how they might influence cyclical processing and connection making.

The Kintsch and van Dijk (1978) model was the basis for a computational model developed by Miller and Kintsch (1980). The Miller and Kintsch (1980) model consisted of two components: a chunking program and a microstructure coherence program. In limiting their model to these two components, the Miller and Kintsch model focused on the subset of the Kintsch and van Dijk processing theory that was concerned with local (cycle to cycle, proposition to proposition) coherence and strategies for resolving breaks in local coherence.

The *chunking program* operated by reading one word at a time from the text, identifying the proposition or propositions associated with the word, and then deciding whether or not the current proposition under consideration should be added to the current "chunk of propositions." The minimum number of words per chunk was specified by the *input size parameter* ***I***.

The *microstructure coherence program* operated by processing a chunk of propositions on each processing cycle. One proposition was designated to be the superordinate proposition and placed at the top of a hierarchical *working memory coherence graph.* The designation of the superordinate proposition had to be done "outside" the computational model by a human modeler and relied on often sophisticated use of that individual's prior knowledge. In the working-memory coherence graph, propositions that were semantically similar to the superordinate proposition were located at levels "higher" in the hierarchy. Semantic similarity was determined by the presence of overlap in the arguments (nouns) in the propositions. The *buffer-size parameter* ***s*** determined the number of propositions that were kept active during processing of the next cycle of propositions. When the number of propositions in working memory at the end of a cycle was greater than ***s***, priority for being held over was based on level in the coherence graph hierarchy and recency of processing, a form of the Kintsch and van Dijk (1978) "leading edge" strategy.

In addition to the working-memory coherence graph, a *long-term memory coherence g*raph was constructed. It differed from the working memory graph in that all propositions that were processed on any cycle were represented. When a new input could not be connected to a proposition in the working memory graph, a reinstatement search of the long-term memory graph ensued. Propositions that provided links to "dangling" new propositions were incorporated back into the current working-memory coherence graph. Failure of the reinstatement search to provide a linking proposition resulted in a coherence break. Coherence breaks were remedied by making inferences that brought in information from prior knowledge but this process was not part of the computational model developed by Miller and Kintsch (1980). What was computationally modeled was that when a coherence break occurred, a new working-memory coherence graph was created, with a new superordinate node (determined outside of the computational model). The probability that a proposition would be recalled was computed by the formula $1-(1-\boldsymbol{p})^{\boldsymbol{n}}$ where $\boldsymbol{n}$ is the number of processing cycles a proposition was maintained in the working memory buffer and $\boldsymbol{p}$ is a free parameter that corresponds to the probability that a proposition will be recalled if it was entered in only one ($n=1$) processing cycle in working memory. Thus, $\boldsymbol{p}$ is a “base” recall probability and all propositions start out at that level once they are processed in a cycle. With each additional working memory cycle in which a proposition is processed, the likelihood of recall increased.

Miller and Kintsch (1980) evaluated the computational model against data from 120 participants who had each read and recalled 20 paragraph-length texts of varying complexity. They used the computational model to predict the expected recall frequency of particular propositions in each of the 20 texts. They found a positive correlation of .6 between observed and predicted recall frequencies. This correlation was considerably less than Kintsch and van Dijk (1978) and other researchers (e.g., Spilich, et al. 1979) had obtained when they had constructed working- and long-term memory coherence graphs by hand and had tested alternative values of the parameter $\boldsymbol{s}$. Miller and Kintsch (1980) attributed the lower performance of their model to the lack of a component that generated macropropositions and the macrostructure of the text that resulted from applying the hierarchical organization rules that applied to micropropositions. Thus, the apparent limitations of a computational model that did not exploit macropropositions led to further development of this aspect of the text comprehension theory, treated at length in the 1983 book *Strategies of Discourse Comprehension* (van Dijk & Kintsch, 1983).

The text comprehension theory detailed in van Dijk & Kintsch (1983) devoted a large amount of attention to how comprehenders translate lengthy, detailed texts into more summary-like representations that rely on frequent and judicious application of rules for substituting a single macroproposition for groups of micropropositions. It also developed the theory of representations and moved from explicit and implicit textbase to a three-level theory of representation. Specifically, van Dijk and Kintsch (1983) postulated that mental representations of text had multiple layers that captured different aspects of text, including the surface form (the specific words, sentences, layout of the text), the meaning of the text itself (textbase), and the interpretation or model of the world referred to by the text (mental or situation model) (van Dijk & Kintsch, 1983). The textbase captures the referential and intra- and inter-sentential relations among the words in the text. The textbase representation maps most clearly onto the earlier local coherence graph. The situation model reflects the integration of prior knowledge with the information explicitly “in” the text. The claim was that situation model construction increased the likelihood that the information could be used in new situations. There were a number of behavioral demonstrations of the validity of both the layers of representations and the importance of macroproposition and macrostructure creation (e.g., Fletcher, 1994; E. Kintsch, 1990; Kintsch et al., 1993; McNamara, Kintsch, E., Songer, & Kintsch, W. 1996; Perrig & Kintsch, 1985; Schmalhofer & Glavanov, 1986). However, efforts to formalize the 1983 version of the theory and develop computational models of it proved elusive, due in large measure to the importance of strategic and prior knowledge in generating macropropositions and situation models. There seemed no *a priori* computational techniques suitable for modeling the strategic management of prior knowledge.

In the face of the computational intractability of the 1983 text comprehension theory, Kintsch proposed a radically different form of text comprehension theory (Kintsch, 1988) that "managed" prior knowledge through non-strategic, associative processes. The model, called the Construction-Integration (CI) model (Kintsch, 1988), is another interesting example of how the *failure* to formulate a computational model of the theory provided impetus for the formulation of a radically different theoretical proposal.

Conceptually, CI is a two-phase, constraint-satisfaction process model (Kintsch, 1988; 1998) in which there is no reliance on strategic processing mechanisms and macrostructure construction. Kintsch (1988) described it as a "dumb" model. The *construction* phase is a text-based, bottom-up process that results in an initial and frequently incoherent representation of the concepts and ideas in the text plus those elements of prior knowledge that are activated by the concepts and ideas/propositions from the text. Concepts and propositions are represented by nodes in a semantic-network like representation. Links among nodes reflect sentence and text-level semantic and logical connections among the nodes. During the *integration* phase, activation is distributed among the nodes and links according to a connectionist algorithm that has the effect of strengthening the nodes that have a lot of connections and are therefore central to the meaning and situation and neglecting those with few connections. Nodes with few connections are often associates to an individual concept but irrelevant to the meaning in the context of the developing network, or are inconsistent with the core meaning. In effect, concepts and ideas that are compatible mutually enhance one another and ones that are incompatible or irrelevant are "ignored." Thus, during integration relevant knowledge becomes more strongly connected to ideas from the text and gaps among ideas are filled in with prior knowledge that is activated through associative memory processes that are consistent with contemporary theories of memory storage and retrieval (e.g., Diller, Noble, & Shiffrin, 2001; Gillund & Shiffrin, 1984; Hintzman, 1988; McClelland & Rumelhart, 1985; Murdock, 1982; Raajmakers & Shiffrin, 1981; Shiffrin & Steyvers, 1997).

Computationally, Kintsch (1988; 1998) modeled CI as a connectionist network (e.g., Rumelhart, Hinton, & McClelland, 1986) of nodes and links among them, arrayed as a matrix in which nodes are the row and column headers and non-zero entries in the cells of the matrix indicate a relation or link between the header nodes for that cell. Each node and link has associated with it an initial, numerical *activation* value. The construction phase builds the matrix and fills in the non-zero cell values, resulting in the *coherence* matrix (Kintsch, 1988). The integration phase then takes over and iteratively applies an activation updating rule. Specifically, all nodes are typically initially activated and then each node updates its activation by computing a weighted sum of the links entering the node and the activation levels of the other nodes in the network attached to the node via those links. All nodes in the network simultaneously update their activation levels and then the activation levels of all nodes is reduced by a fixed amount to prevent activation levels from growing without bounds. Under general conditions, the activations when updated in this manner will eventually tend to stop changing (Guha & Rossi, 2001).

When the change in activation levels across the nodes becomes minimal across iterations, the integration phase ends and the resulting activation values of the nodes and links are “saved” in a long term memory matrix of connection strengths using a version of the Hebbian learning rule as described in Kintsch (1988). These connection strengths (or equivalently “links”) among nodes are additively updated if the link participates in additional processing cycles. Typically, when a sentence is processed, it produces a matrix in which each noun is a concept node and the verb generates a predicate proposition node that references the concept nodes. Thus the cells in the matrix capture the intersection of the concept nodes with themselves and the intersection of the concept nodes with the predicate proposition (when drawing a network representation, nonzero entries in the cells correspond to links between nodes). Assuming a constant activation parameter, the predicate proposition receives greater initial activation than the concept nodes it relates. When successive sentences in a text are processed, the CI model adopts the assumption of cycles of input in a limited working memory environment that was part of the 1978 and 1983 versions of Kintsch and van Dijk’s theory. Across cycles of the construction process, links between predicate propositions are formed if they are present during the same construction cycle and if there is overlap between them. A frequently made assumption is that the nodes most active at the end of a cycle are carried into the next input cycle (Kintsch, 1988; 1998). Although 2 is the number frequently used for this “carry over” parameter, modelers have manipulated this value, sometimes finding better fits of the model for larger values and sometimes not (e.g., Tapiero & Denhière, 1995).

Simulations of behavioral data based on the CI model (Kintsch, 1988, 1998) have resulted in moderate to good correlations between the model's performance and human performance across a range of comprehension and learning tasks (Kintsch, 1998; Kintsch & Greeno, 1985; Singer & Kintsch, 2001; Wolfe & Goldman, 2003), although the predictions have typically been better for memory tasks than for on-line processing tasks. Furthermore, in implementing the CI model, there are - quite understandably - many places where modelers must make decisions about various parameters (e.g., the number of propositions to bring in on a cycle; the number and which propositions to carry over to the next cycle; initial activation values, weighting of different kinds of relationships among nodes; how much and what prior knowledge to include in the construction phase; what the relations are among nodes across textbase and situation model levels of the representation, and so on). As a result, the CI model has prompted the development of a number of additional computational models that bear a family resemblance to CI but that make different assumptions about one or more of the components or parameters of the computational processing model, including the operation of working memory and the carry-over parameter (Goldman & Varma, 1995; Goldman, Varma, & Coté, 1996; Langston & Trabasso, 1999; Tapiero & Denhière,1995), the learning algorithm (Goldman, et al., 1996; van den Broek, Risden, Fletcher, & Thurlow, 1996; van den Broek, Young, Tzeng, & Linderholm, 1998), and the basis of establishing connections among nodes in the coherence matrix (Langston & Trabasso, 1999; van den Broek et al., 1998).

All of these computational modeling efforts have helped define important yet unresolved issues in text comprehension or have presented convincing evidence for the utility of the particular computational model that was tested. In so doing these computational modeling efforts have spurred the development of text comprehension theory. In the present context we highlight two modeling efforts in the CI family. The first, a relatively close relative to CI, is the Capacity-Constrained Construction Integration (3CI) model (Goldman & Varma, 1995; Goldman et al., 1996). It examined an alternative conception of working memory processes but otherwise remained faithful to the assumptions of the CI model. The second case, Landscape theory, (van den Broek et al, 1998; van den Broek et al., 1996) is a more distant cousin to CI and makes different assumptions about a number of process mechanisms.

The Capacity-Constrained Construction – Integration (3CI) Model

The 3CI model altered the working memory mechanism of the CI model. Goldman and Varma (1995; Goldman et al., 1996) used the computational architecture of the Just and Carpenter Collaborative Activation-base Production system model (3CAPS) so they could substitute a dynamic working memory process for the fixed working memory parameter ***s*** in the CI model. The critical feature of 3CAPS for the Goldman and Varma (1995) 3CI model is the assumption that elements active in working memory compete with one another for activation in a limited or capacity-constrained working memory. Elements gain and lose activation dynamically. Processing in the 3CI model operates on a cycle to cycle basis. The more activation an element starts a processing cycle with, the more likely it is to accrue activation on that cycle, as is true in the CI model. Different from the CI model is that there is no forced removal (or decision to “hold over” some propositions and delete others) of specific propositions for processing with the next cycle of input. Rather, as elements decrease in activation, they become less available for connection with other elements, eventually falling to such low levels that they are effectively no longer “present” in working memory. As in the CI model, at the conclusion of each processing cycle activation levels of elements and links among them are updated in a long term memory matrix. Strengths in this matrix are the basis for predicting the likelihood of inclusion in recall.

Goldman and Varma (1995) applied the 3CI and the CI model to the same sets of recall data to examine the differences in the predictions made by the alternative models. The behavioral data had been obtained from adults and from children who had read short, informational passages (250-300 word) that had a hierarchical global structure. The 3CI model produced a pattern of activation levels across the passage sentences that mimicked the global structure of the passage whereas the CI model produced activation patterns that were sensitive only to the local, sentence-to-sentence structure. That is, the 3CI model produced higher activations for topic sentences relative to the detail sentences of each paragraph in the passage, corresponding to the hierarchical content structure of the passage. Recall predictions that were derived from the 3CI model significantly correlated with behavioral recall data from adults. As a group, the adults recall patterns showed sensitivity to the global structure of the passage in that they recalled main ideas more frequently than the details that elaborated them. However, among the children the distinction between main ideas and details was far less obvious. The two models were equally good at predicting the children’s data when the students did written recall. When children orally recalled what they had read, CI correlated with recall performance better than 3CI did.

The comparative predictive ability of 3CI versus CI was tested further by examining the ability of each model to predict recall for informational passages that had different content structures from the passages examined in the initial comparison (Goldman, et al., 1996). Two findings are particularly relevant to the current point. A detailed analysis of the sentences for which the computational models underpredicted (i.e., the sentence was recalled more frequently than predicted by the computational models) behavioral recall data indicated that these tended to be of two types. First, sentences in which the information was highly familiar to readers, e.g., *Dentists have to fix cavities,* were underpredicted by both 3CI and CI. This is understandable because neither model had been implemented with a mechanism for incorporating prior knowledge into the construction process. This underprediction led Goldman and colleagues to argue for the need to include situation-model nodes as well as a principled means of introducing prior knowledge into the construction and integration phases of text processing.

The other kind of sentence that was underpredicted were those that had high overlap with prior sentences and were important to the content structure but that came late in the passage. The underprediction of these turns out to be the result of a property of an evolving network of propositions in which new and old propositions compete for available activation. As the network gets larger and more stable, it essentially feeds itself and it is more difficult for a new proposition to accrue sufficient activation to “break into” the network. To deal with this problem Goldman and colleagues modified the way the integration process operated in the 3CI model. They incorporated a “top end” activation threshold: Once a proposition exceeded this threshold, it no longer competed with other propositions for activation, allowing new input to have a greater chance of accruing activation. This threshold cap embodied the notion that some ideas are so prevalent in a passage that once they reach a certain strength (the threshold) they will be remembered regardless of what else comes in. The top end threshold essentially substituted a sigmoidal for a linear activation function.

Interestingly, neither CI nor 3CI were able to account for the online processing of the passages. Reading time data were not predicted by the number of cycles needed for the network to settle, a measure derived from the integration phase of the modeling. Thus, the 3CI effort advanced the theory in the area of working memory processes but did not shed light on the predictions of reading time. This should not be surprising because a large amount of the variance in reading time is predicted by many characteristics of the surface, input text (Haberlandt & Graesser, 1985). Both CI and 3CI operate on propositional input rather than surface text sentences. The lack of prediction of processing time and its relation to characteristics of the surface text of passages underlines the importance of developing ways to parse the input language of the text, a significant computational challenge. New theories of parsing and syntactic analysis are emerging, however. Some of these appear promising for use in computational models of text comprehension (e.g., Dennis, 2004; Durbin et al., 2000). For example, Golden and his colleagues (Durbin et al., 2000; Ghiasinejad & Golden, submitted) are developing a computational model for automatically identifying the presence of propositions in free response data. The essential idea of the computational model is that representative free response data are first semantically annotated using a semantic annotation system embodied within a user-friendly software interface. The computational model then learns statistical regularities between subsequences of words in the free response data and the semantic annotations by interacting with an experienced human coder. Specifically, the percentage of times that a word is used to express a particular word-concept and the percentage of times that one word-concept follows another when expressing a particular proposition is recorded during the learning process. Eventually the system (in relatively constrained task domains) is capable of automatic identification of propositions in free response data.

The Landscape Theory and Computational Model

The second example we elaborate is the Landscape theory, developed by van den Broek and colleagues (van den Broek et al., 1998; van den Broek et al. 1996). As indicated in a prior section of the chapter, Landscape theory shares features with CI theory but differs in several important ways. First, Landscape theory posits a dynamic and reciprocal interaction between online processes and the gradually emerging offline product of reading. Second, readers' goals and judgments of coherence are integral to the architecture of the Landscape model and their relation to reading processes is explicit. Third, Landscape theory treats coherence as arising from multiple representational dimensions and their interactions (see also Zwaan, Magliano, & Graesser, 1995) and connects these dimensions to the readers' standards of coherence in that situation. This contrasts with other theories that typically focus on a single dimension of coherence.

The Landscape theory captures both on-line comprehension processes and memory performance after reading is completed. In this theory reading is conceived as a cyclical process, in which propositions (or other units of text) fluctuate in their activation from one cycle to the next. There are several major sources of activation at each cycle: the current input cycle, the preceding cycle (through carry over), the memory representation of the text as constructed in the preceding processing cycles, and background knowledge. The last two –memory for the text read so far and background knowledge- can be accessed through a spread of activation process (called *cohort activation)* or through strategic (re)instatement. Together with working memory or attentional limitations, these sources result in an *activation vector* that forms the basis for updating the episodic memory representation for the text. In the computational implementation of the Landscape model, at each cycle the representational node strength of a proposition increases as a function of the amount of activation it receives. In addition, a connection is established (or, in the case of an existing connection, strengthened) between co-activated propositions, as a function of the amount of activation each receives. A central component of the computational model is that the activations vectors and the developing memory representation interact dynamically: with each reading cycle the memory representation is updated and, in turn, the updated memory representation strongly influences subsequent activation vectors. Another central component of Landscape theory is that in each reading situation a reader applies a particular set of *standards of coherence* (van den Broek, Risden & Husebye-Hartman, 1995; see also Goldman, et al., 1996). At each individual reading cycle, these standards determine whether the information activated through cohort activation is adequate to satisfy the reader or whether strategic processes are required. Standards of coherence differ across readers and across reading situations, depending on reading goal, task demands, textual properties, and so on, but in most cases they include at least standards of referential and causal coherence. From a computational standpoint, standards of coherence set a threshold value. If threshold is met or exceeded, the reader proceeds to the next input cycle; otherwise, processing of the current cycle continues.

To keep track of the many components and their interactions, van den Broek and colleagues implemented the theory in a computational model (van den Broek et al., 1998; van den Broek et al., 1996; Linderholm,Virtue, van den Broek, & Tzeng, 2004). Tests of the computational model showed that it did a good job predicting behavioral data. The model's predictions for on-line activations and frequency of off line recall correlated between .55 and .65 with readers' data. Furthermore the change in activation vector from one cycle to the next (called the *activation gradient*) predicted reading times for the second cycle. As a final example, the model does a good job postdicting the inconsistency detection data reported by O'Brien and Albrecht (1992) and the effects of reading goal on inference generation reported by van den Broek, Lorch, Linderholm, and Gustafson (2001).

Development of the computational model allowed initial testing of Landscape theory and showed that it captures a wide array of phenomena observed in the reading process and representation construction. Equally important for our current purpose however is that the process of creating the computational model led to considerable development of Landscape theory. For example, to implement computationally the notion that the activation vectors result in (or update an existing) memory representation it was necessary to provide an explicit 'mini' theory of exactly *how* such construction/updating occurs. Such a 'mini' theory had to specify the precise manner in which co-activation leads to connection construction:

- Is the connection strength that results from an activation vector all-or-none (i.e., if two propositions are co-activated a connection is forged regardless of their actual activation values), additive (i.e., the connection strength is the sum or each of their activations or if one allows negative activations, the sum of the absolute values of their activations), or multiplicative (i.e., the connections strength is a function of the product of the two activations)?
- Do subsequent co-activations change the strength of an existing connection in a linear fashion or in a non-linear (e.g., asymptotic) fashion?

Findings in prior research in memory and in connectionist models formed the basis for a theoretical component that made the translation from activation vector to memory representation explicit. With regard to the examples above, the mini-theory assumes that the change in connection strength is a multiplicative function and that updating follows an asymptotic curve.

A second contribution to the development of theory concerns the fact that both the episodic memory representation and semantic background knowledge are presumed to be accessed via cohort activation. When it came to deciding on parameters to describe such spread of activation, a choice had to be made whether the parameter settings would be identical for the sources of activation. By allowing the parameters to differ, it is possible to consider differential ‘weights’ for the two sources. Thus, the translation of the Landscape theory into a computational model stimulated the development of further theoretical notions as well as the precise specification of the existing theory.

Summary

The evolution of Kintsch's comprehension theory along with examples of additional computational models of text comprehension have been used to illustrate ways in which building computational models from theoretical formulations of text comprehension has resulted in advances in text comprehension theory. In the process we have reported some of the behavioral data that modelers have attempted to explain. Computational models are often used to help formalize relationships in behavior data. In the process they are sometimes able to help make sense of both expected and unexpected patterns in behavioral data.

Computational Models Assist in Making Sense of Surprising Behavioral Data

In addition to providing impetus for the development of theory, computational modeling can help provide and/or test post hoc explanations of behavioral data whose patterns differ from *a priori* predictions, are surprising, or seem contradictory. For example, the features of the computational Landscape model led to unexpected –and theoretically important- predictions. For example, by adding input cycles that were 'empty' (i.e., zero-vector that did not contain activation for any propositions) the patterns of activation in the final activation vector and the connection matrix that constitutes the final memory representation were altered in structural ways. By comparing the two sets of predictions (before and after the empty cycles) to human data, van den Broek and colleagues noticed that the first set (before empty cycles) predicted immediate recall well but was much poorer predicting delayed recall; the pattern for the second set (after empty cycles were added) was the reverse: much better predictions for delayed than for immediate recall. These observations suggested that a major difference between immediate and delayed recall consists of a period of no new activation and thereby of additional weeding out of transient activations from the activation vectors as well as memory representation. In addition, they suggested that immediate recall is a function of both the memory representation and the activation vector for the final reading cycle, whereas delayed recall is determined just by the (now further updated) memory representation. These findings and speculations allowed the two types of memory to be included in a single architecture.

A second example involving the Landscape model pertains to the adoption, described above, of an asymptotic learning curve. In the architecture of the Landscape model the asymptotic learning curve resulted in the prediction that the memory representation connection from proposition A to proposition B might differ in strength from proposition B to proposition A. In other words, the connections between propositions were predicted to be *asymmetric*. Although this predictive effect was unintended, similar effects have been extensively documented in the research literature on semantic memory and, on occasion, in the literature on discourse processing (Lutz & Radvansky, 1997; Trabasso, Secco & van den Broek, 1984)

Applications of computational modeling to inferences

Inference making is one area of text processing research that has generated conflicting theories, models, behavioral data, and attendant controversy (c.f., Graesser et al., 1994; McKoon & Ratcliff, 1992; 1995). Computational modeling of various inference tasks and behavioral data is leading to better understanding of some of the issues. In this section we discuss two of these applications, one dealing with the time course of recognition and retrieval memory for inferences as compared to explicitly presented text (Singer & Kintsch, 2001) and the second dealing with different types of inferences (Schmalhofer, McDaniel, & Keefe, 2002).

Recognition and retrieval of inferences

Singer and Kintsch (2001) combined the C-I framework with the Gillund and Shiffrin global memory-matching model (Gillund & Shiffrin, 1984) in an effort to account for a complex pattern in inference memory data collected by Zimny (1987; also reported in Kintsch, Welsch, Schmalhofer, & Zimny, 1990). Zimny tested memory for probe words at three delays using a recognition memory task and a sentence verification task. Probe words were related either to explicit, paraphrased, or inferred text information or were related to distractors. Of particular interest here are the different patterns that Zimny (1987) reported in the two tasks. In both tasks and across the three delay conditions, probes related to explicit information were always recognized the best and at high levels (70 to 80% in recognition; 85 – 95% in verification). However, the pattern for inferences was different depending on the task. In the recognition task, memory for probe words related to inferences grew stronger over time: At immediate test, only 20% of the participants said the probe word had been presented (a false recognition) whereas at the long delay almost 60% said it had been presented. In contrast, in the sentence verification task, participants verified that probe words related to inferences had occurred in the text as often as they said that probes related to explicit information had been presented in the text. In other words, in the sentence verification task, memory for inference-related words was as strong as memory for explicit information at each delay, whereas on the recognition task, memory for inference-related words became stronger over time.

Singer and Kintsch (2001; also see Kintsch et al., 1990) found that they could account for this complex pattern of results by using a version of the CI model to characterize the dynamical changes to the reader's working memory connection matrix by specifying how sentence nodes, proposition nodes, and macrostructure proposition nodes are interconnected. Singer and Kintsch (2001) then used a modified version of the Gillund and Shiffrin (1984) theory of recognition memory to make predictions regarding performance on sentence recognition memory and sentence verification tasks. Specifically, they calculated familiarity of a probe based on its connection strength over the whole coherence matrix, consistent with the global memory-match retrieval mechanism of Gillund and Shiffrin (1984). Furthermore, familiarity calculations were done using the Gillund and Shiffrin (1984) multiplicative combining rule rather than a linear one. Finally, they used response decision rules from signal detection theory and determined different response thresholds for recognition memory and sentence verification tasks. This three-part process produced simulation data that was consistent with the previously obtained behavioral data. Singer and Kintsch (2001) noted that the three parts of the simulation needed to operate together to produce the particular observed qualitative (and quantitative) pattern of predictions. Each was necessary but not sufficient; all three were essential for the model to make the correct qualitative pattern of predictions. Space does not permit us to treat this model in all of its complexity and detail and the interested reader is referred to Singer and Kintsch (2001) for a full explication of the derivation and arguments for it.

Simulating bridging and predictive inferences.

Schmalhofer et al. (2002) pointed out that there are a number of explanations and theories about when and why different kinds of inferences are made. For example, they indicate that inferences that fill in gaps among bits of information that have already been processed (backward or bridging inferences) are made with high probability. In contrast, inferences that predict what will happen next (forward inferences) are made with much less frequency. Their goal was to use the CI model to provide a unifying account of both kinds of inferences. Using materials that Keefe and McDaniel (1993) used to examine bridging and forward inferences, Schmalhofer et al. (2002) constructed connectivity matrices for three levels of representation (surface text, propositional, and situational) as well as the connectivity between levels. Thus, concepts explicitly presented in the text would have multiple levels at which to accrue activation whereas nodes generated from prior knowledge would be represented at the situation level and perhaps at the propositional level. Key to understanding their argument is that nodes accrue activation on the basis of within and between level connectivity. As processing proceeds, new input provides reason to continue to activate specific nodes at all levels; if there is not input that connects to an activated node on a particular cycle it will lose activation and be less likely to show priming effects.

Connectivity is the unifying principle in the Schmalhofer account. Nodes that are more highly connected to other nodes, especially if this is sustained over multiple processing cycles, would show an increased likelihood of a priming effect, regardless of whether the node represents a bridging or predictive inference. CI simulations based on the derivation of the connectivity matrices, paying close attention to within and across level connectivity, yielded both qualitative and quantitative predictions that were consistent with the data of Keefe and McDaniel (1993). Schmalhofer and colleagues were thus able to account for the time-course of activation of predictive and bridging inferences.

The use of computational models to account for patterns of behavioral data, expected as well as unexpected, helps to integrate and unify empirical findings and support theory development. In developing these models, researchers are forced to be quite explicit regarding the mechanisms, processes, and relations among them. This characteristic of computational models enables better communication.

Computational Models Support Communication

The precise specification required to enact computational models facilitates communication among researchers working in similar areas as well as those working in seemingly unrelated areas. First, we discuss the issue of automated coding of free response data. Second, we discuss how computational models are useful for integrating comprehension and memory, areas that are typically seen as quite related. Our final example illustrates the communication role of computational models through the use of a text comprehension model to account for decision making data.

Reliable and documented coding of free response data

A typical procedure for coding protocol data involves having two experienced human coders work together in the analysis of a portion of the protocol data. Critical propositions and methodologies for identifying such propositions in an objective manner as possible are then developed by the coders. The remaining portion of the protocol data is then coded independently by the two coders for the purposes of computing a measure of inter-coder reliability. Typically, in text comprehension research, agreement measures in the 95% range with Cohen Kappa (Cohen, 1960; Carletta, 1966) scores in the 70% range are considered to establish acceptable and reliable coding procedures.

This widely used methodology for coding verbal protocol data, however, suffers from a variety of serious intrinsic problems. First, despite the best of efforts, explicit details governing all aspects of how free response data is mapped into a propositional representation can never be provided by the above procedure. There will always be a subjective component to the above process. Second, even if all details of the coding procedure could be explicitly documented, there is no guarantee that the resulting coding procedure would always be applied in a consistent manner by human coders. Third, efficient unambiguous communication of complex ideas is an essential component of science. Even if all details of the coding process could be explicitly documented and then always consistently implemented without error, the resulting coding process (as typically implemented in the current scientific literature) would probably be highly complex and difficult to efficiently communicate to other scientists. If the efficiency of such communications could be improved, then the measurement of detailed methodological coding issues upon experimental behavioral findings would be facilitated. In addition, replications of experimental findings across research labs could be improved as well. And finally, detailed semantic coding of protocol data tends to be time consuming and effort-intensive. If the costs of data analysis could be reduced, then protocol data could be analyzed more rapidly which would ultimately increase the overall rate of scientific progress in the area of discourse processes. Ericsson and Simon (1984) provide a further discussion of these issues.

These concerns suggest that an important challenge for text comprehension research in the next century will be to aggressively incorporate tools from artificial intelligence to facilitate the automatic coding (or at least support manual coding) of free response data. Some important steps in this direction have already been taken but more work needs to be done. Examples of progress in this area include: the string theory approach of Dennis (2004; also see Dennis [this volume]), the Hidden Markov Model approach of Ghiasinejad and Golden (submitted; also see Durbin, Earwood, and Golden, 2000), and Latent Semantic Analysis methodologies (Dunn et al., 2002; Foltz et al., 1998; Landauer & Dumais, 1997), and the probabilistic automated semantic role labeling methodology of Gildea and Jurafsky (2002).

Integrating text comprehension and memory

Earlier in the chapter we discussed examples of the integration of memory models with text comprehension theories and resulting improvements in model fits to behavioral data (e.g., Goldman & Varma, 1995; Singer & Kintsch, 2001; Schmalhofer, et al., 2002). Other examples include the work of Fletcher, van den Broek, and Arthur (1996) who used an alternative modification of the Gillund and Shiffrin (1984) model to develop a theory of text recall based upon local coherence strategies. Fletcher et al. (1996) found that the resulting model provided good predictions of what propositions were recalled by participants as well as the order in which the propositions were recalled.

There are also efforts to integrate models of prior knowledge with text comprehension. Kintsch (1998) has incorporated Latent Semantic Analysis (LSA; Landauer & Dumais, 1997) as the engine for generating prior knowledge elements during the construction phase of CI. Briefly, LSA is a computational approach to word meaning that is based on co-occurrences of words in printed text from which semantic spaces that reflect meaning relationships among words are derived. In the context of text comprehension, concepts are added to the situation level representation based on their similarity of meaning with the words in the text. The integration process then operates and only those nodes that are relevant in the context tend to receive higher activation and become part of the situation level representation. In a further elaboration of the use of LSA in comprehension, Kintsch (2001) has proposed a predication model that enables computational modeling of metaphor comprehension.

In a somewhat different vein, computational modeling has been useful in efforts to understand the resonance theory account of the "distance effect" (Myers & O'Brien, 1998). O'Brien, Plewes, & Albrecht (1990) identified the distance effect in their research on situations involving two potential referential antecedents for a referent in an incoming target sentence. The referential antecedents were positioned in the text so that they would normally not be strongly activated in working memory when the target sentence was read. Given this situation, O'Brien et al., (1990) found that the referential antecedent that is closest to the target sentence will be more strongly activated in working memory relative to the antecedent that is furthest from the target sentence. Myers and O'Brien (1998) interpreted these results as supporting a resonance memory theory that asserts that information in both working memory and long-term memory is available for re-activation in working memory given appropriate retrieval cues.

However, based on previously published accounts of the resonance model, Lutz and Radvansky (1997) concluded that the resonance model would always predict a “distance effect” (i.e., increasing the amount of intervening text between a referential or causal antecedent and its target sentence would tend to decrease the activation of the antecedent in working memory when the target sentence was processed). Myers and O’Brien (1998), however, emphasized that in the resonance model the presence or absence of a distance effect is not merely a function of the number of intervening statements between the antecedent but is also a complex function of the propositional content of the text passage. To illustrate their point, they proposed a two-parameter model that could simultaneously capture the presence of a distance effect for experimental texts from studies in which the behavioral effect was observed, as well as the absence of such a distance effect for texts where no distance effect was observed in the behavioral data (e.g., Lutz & Radvansky (1997). By expressing their theory as a computational model, Myers and O’Brien (1998) provided a medium for the communication and evaluation of the structural properties and implications of a particular explicit model of reading comprehension processes.

Integrating text comprehension mechanisms with decision making

Support for a number of phenomena in social psychology and decision making rely on the use of vignettes or short texts about people and situations in which they find themselves. For example, Kahneman and Tversky (1982) had participants read a story about a Mr. Jones who left his office, did not drive home by his regular route, stopped at a light, and then got killed by a speeding truck at an intersection. According to “norm theory” in the decision making literature (Kahneman & Miller, 1986), it is easier for decision-makers to construct typical alternatives to typical events rather than atypical alternatives to typical events. Thus, norm theory would predict that decision makers (upon encountering this story) have a tendency to focus upon the unusual causal antecedent since the statement of the unusual event tends to evoke normal alternatives. Thus, the unusual antecedent is usually viewed by participants reading the text as the reason for the traffic accident.

Trabasso and Bartolone (2003) provided an explanation for the “unusual antecedent” phenomenon based on text comprehension processes and the use of the resulting representation to make decisions about possible causes of the accident. They used the discourse analysis techniques of Trabasso, Secco, and van den Broek (1984) to create a causal network of clauses and causal links among them. Integration of the network occurs via a connectionist model (Langston & Trabasso, 1999) to produce connection strengths for the various clauses. The connection strengths index accessibility of various clauses as explanations for specific events. Trabasso and Bartolone’s (2003) analysis showed that in the story with the unusual route, there were more events explaining why that route had been taken than there were explanatory events for the typical route in the typical-route story. They hypothesized that the explanatory focus on the unusual event might make it more accessible as a cause for the accident. Trabasso and Bartolone (2003) tested this hypothesis by constructing a series of variations of the Mr. Jones text that systematically and independently manipulated the typicality and explanation variables. They constructed the causal networks for each and integrated them using their connectionist simulation model. The causal network construction showed different patterns of connections for the different versions of the stories and hence differential connection strength and accessibility values resulting from integration using the Langston and Trabasso model (1999; also see Langston, Trabasso, and Magliano, 1998). Indeed, the connection strengths that resulted from the simulations showed that explanation but not typicality was the essential variable in readers' decisions about what caused the accident.

The discourse analysis and computational modeling conducted by Trabasso and Bartolone (2003) shows that causal explanation plays a powerful role in both text comprehension and decision making. Both the detailed analysis of the connections among the events and the computational modeling of those connections in terms of strength and accessibility were necessary to make a forceful and convincing argument regarding the centrality of explanation in both comprehension and decision making.

Summary and Conclusions

We have provided three answers to the question *Why are computational models of text comprehension useful?* First, we illustrated the role of computational models in the evolution of theories of text comprehension. Both success and failures of computational models were shown to be informative for theory development. Second, the computational models were shown to be useful for testing explanatory constructs and accounting for unexpected findings. The use of computational models allows for the explication of the mechanisms involved in performing specific text comprehension tasks. Creating models of these mechanisms and tasks that are then successful at replicating qualitative patterns in behavioral data adds plausibility to explanatory constructs and may shed light on unifying constructs. Third, we provided a discussion of the role of computational models in enabling and enhancing communication with and across areas of work in psychology. Although we used different examples to illustrate each of these contributions of computational models, all computational models have the potential for increasing communication and most contribute, albeit indirectly, to theory development.

Our examples were necessarily limited, drawing on just a subset of the computational models that have contributed to the advancement of text comprehension. The other chapters in this volume bring to the reader cutting-edge work on new and emerging computational approaches that are increasingly multidisciplinary. In a multidisciplinary context precise communication is even more important than among researchers from the same discipline. As researchers in the fields of computer science, neuroscience, and cognitive psychology attempt to reconcile their findings and theories, the communicative value of computational models takes on even greater importance than it has had until now. Specification sufficient for computational modeling will serve to clarify the intentions of the models and make the outcomes and implications easier to evaluate and interpret. With this level of clarity, multidisciplinary discussions can benefit from and build on the cumulative knowledge base resulting from theoretical and empirical advancements in text comprehension research.

References

Carletta, J. (1996). Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics*, *22*, 250-254.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurements*, 20, 20, 37-46.

Dennis, S. (2004). An unsupervised method for the extraction of propositional information from text. *Proceedings of the National Academy of Sciences, 101*, 5206 – 5213.

Dennis (???) Book chapter in this volume??? Delete if there isn't one.

Diller, D. E., Nobel, P.A., & Shiffrin, R. M. (2001). An ARC-REM model for accuracy and response time in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 414-435.

Dunn, J. C., Almeida, O. P., Barclay, L., Waterreus, A., & Flicker, L. (2002). Latent semantic analysis: A new method to measure prose recall. *Journal of Clinical and Experimental Neuropsychology*, 24 (1), 26-35.

Durbin, M. A., Earwood, J., & Golden, R. M. (2000). Hidden Markov models for coding story recall data. In *Proceedings of the 22nd Annual Cognitive Science Society Conference*, pg. 113-118. Mahwah, NJ: Erlbaum.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Fletcher, C. R. (1994). Levels of representation in memory for discourse. In M. A. Gernsbacher (Ed), *Handbook of psycholinguistics* (pp. 589-607). San Diego, CA: Academic.

Fletcher, C. R., van den Broek, P., & Arthur, E. J. (1996). A model of narrative comprehension and recall. In B. K. Britton and A. C. Graesser (Eds.), *Models of Understanding Text* (pp. 141-163). Mahwah, NJ: Erlbaum.

Gernsbacher, M. A. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Erlbaum.

Gildea, D., & Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational Linguistics, 28,* 245-288.

Ghiasinejad, S. & Golden, R. M. (submitted). *An empirical evaluation of the AUTOCODER system for automatic semantic coding of children summarization data.*

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1-67.

Goldman, S. R., & Varma, S. (1995). CAPping the construction-integration model of discourse comprehension. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 337-358). Mahwah, NJ: Erlbaum.

Goldman, S. R., Varma, S., & Coté, N. (1996). Extending capacity-constrained construction integration: Toward "smarter" and flexible models of text comprehension. In B. K. Britton & A. C. Graesser (Eds.), *Models of text comprehension* (pp. 73-113). Hillsdale, NJ: Erlbaum.

Goldman, S. & Varnhagen, C. (1986). Memory for embedded and sequential story structures. *Journal of Memory and Language, 25,* 401-418.

Gough, P. (1972). One second of reading. In J. F. Kavanaugh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading* (pp. 331-358).

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 95,* 163-182.

Guha, A, & Rossi, J. P. (2001). Convergence of the integration dynamics of the construction-integration model. *Journal of Mathematical Psychology, 45,* 355-369.

Haberlandt, K., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interations. *Journal of Experimental Psychology: General, 114,*357-374.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528-551.

Just, M. A., & Carpenter, P. A. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Kahneman, D., & Miller, D. T. (1986). *Norm theory: Comparing reality to its alternatives. Psychological Review, 93,* 136-153.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201-208). New York: Cambridge.

Keefe, D. E., & McDaniel, M. A. (1993). The time course and durability of predictive inferences. *Journal of Memory and Language, 32,* 446-463.

Kintsch, E. (1990). Macroprocesses and microprocesses in the development of summarization skill. *Cognition and Instruction, 7,* 161 – 195.

Kintsch, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ: Erlbaum.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95,* 163-182.

Kintsch, W. (1998). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*, 163-182.

Kintsch, W. (2001). Predication. *Cognitive Science, 25*, 173-202.

Kintsch, W., Britton, B. B., Fletcher, C. R., Kintsch, E., Mannes, S. M., and Nathan, M. J. (1993). A comprehension-based approach to learning and understanding. In D. Medin (Ed.), *The psychology of learning and motivation, Vol. 30* (pp. 165-214). NY: Academic Press.

Kintsch, W., & Greeno, J. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*, 109-129.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Kintsch, W., & Welsch, D. M. (1991). The construction-integration model: A framework for studying memory for text. In. W. Hockley & S. Lewandowsky (Eds.), *Relating theory anddata: Essays on human memory in honor of Bennet B. Murdock* (pp. 367-386). Hillsdale, NJ: Erlbaum.

Kintsch, W., Welsch, D. M., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: Theoretical analysis. *Journal of Memory and Language, 29,* 133-159.

Landauer, T.K. and 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.

Langston, M., & Trabasso, T. (1999). Modeling causal integration and availability of information during comprehension of narrative texts. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 29 - 69). Mahwah, NJ: Erlbaum.

Langston, M., Trabasso, T., & Magliano, J. P. (1999). A connectionist model of narrative comprehension. In A. Ram and K. Moorman (Eds.), *Understanding language understanding: Computational models of reading*, (pp.). Cambridge: MA: MIT Press.

Linderholm, T., Virtue, S., van den Broek, P., & Tzeng, Y. (2004). Fluctuations in the availability of information during reading: Capturing cognitive processes using the landscape model. *Discourse Processes.*

Lutz, M. F., & Radvansky, G. A. (1997). The fate of completed goal information in narrative comprehension. *Journal of Memory and Language, 29,* 469-492.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information*. Journal of Experimental Psychology: General, 114,* 159-188.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99,* 440-466.

McKoon, G., & Ratcliff, R. (1995). The minimalist hypothesis: Directions for research. In In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 97-116). Mahwah, NJ: Erlbaum.

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.

Miller, J. R. & Kintsch, W. (1980). Readability and recall of short prose passages: A theoretical analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6,* 335-354.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89,* 609-626.

Myers, J. L. & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes, 26*, 131-157.

O'Brien, E. J. & Albrecht, J. E. (1992). Comprehension strae4gies in the development of a mental model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*,777-784.

O'Brien, E. J., Plewes, S., & Albrecht, J. E. (1990). Antecedent retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 346-352.

Perrig, W. & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language, 24,* 503-519.

Raajmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93-134.

Rumelhart, D.E., Hinton, G.E., & McClelland, J.L. (1986). A general framework for parallel distributed processing. In Rumelhart, D.E., & McClelland, J.L. and the PDP Research Group

(Eds.) *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (pp. 45 – 76) . MIT Press: Cambridge, MA.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a progrmmer's manual: Verbatim, propositional, and situation representations. *Journal of Memory and Language, 25,* 279-294.

Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A unified model of predictive and bridging inferences. *Discourse Processes, 33,* 105-132.

Shiffrin, R. M., & Steyvers, M. (1994). Model for recognition memory: REM – Retrieving effectively from memory. *Psychonomic Bulletin & Review, 4,* 145-166.

Singer, M., & Kintsch, W., (2001). Text retrieval: A theoretical exploration. *Discourse Processes, 31,* 27-59.

Tapiero, I., & Denhière, G. (1995). Simulating recall and recognition by using Kintsch's construction-integration model. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 211-232). Mahwah, NJ: Erlbaum.

Trabasso, T., & Bartolone, J. (2003). Story understanding and counterfactual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 904-923.

Trabasso, T., Secco, T., & van den Broek, P. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83-111).

van den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 423 - 445). Hillsdale, NJ:Erlbaum.

van den Broek, P., Lorch, R.F. Jr, Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory and Cognition, 29,* 1081-1087.

van den Broek, P, Risden, K., Fletcher, C. R., & Thurlow, R. (1996). A "landscape" view of reading: Fluctuating patterns of activations and the construction of a stable memory representation. In. B. K. Britton & A. C. Graesser (Eds.), *Models of Understanding Text* (pp. 165-187). Mahwah, NJ: Erlbaum.

van den Broek, P., Risden, K., & Husebye-Hartman, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch, Jr. & E. J. O'Brien (Eds.), *Sources of Coherence in Text Comprehension* (pp. 353-373). Hillsdale, NJ: Erlbaum.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1998). The landscape model of reading: Inferences and the on-line construction of a memory representation. In H. van Oostendorp & S.R. Goldman (Eds.), The construction of mental representations during reading (pp. 71-98). Mahwah, NJ: Erlbaum.

van Dijk, T. , & Kintsch. W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Wolfe, M.B.W., & Goldman, S.R., (2003). Use of latent semantic analysis for predicting psychological phenomena: Two issues and proposed solutions. *Behavior Research Methods, Instruments, & Computers, 35(1),* 22-32.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and nonnested hypotheses. *Econometrica, 57,* 307-333.

Zimny, S. T. (1987). *Recognition memory for sentence from a discourse.* Unpublished doctoral dissertation, University of Colorado, Boulder.

Zwaan, R.A., Magliano, J.P., & Graesser, A.C. (1995) Dimensions of situation-model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 386-397.