Causal network analysis validation using synthetic recall protocols

RICHARD M. GOLDEN

University of Texas at Dallas, Richardson, Texas

Subjects read and recalled 12 short texts in a memory recall experiment. The order in which subjects recalled the propositions in the text was recorded. A causal network analysis of each text was then done in order to determine how the propositions in each text were causally related. In addition, an episodic memory network analysis of each text was done in order to represent the original order of propositions presented to each subject in the experiment. The human text recall data were then analyzed using a new statistical methodology known as the *temporal Markov field* (TMF) approach, which makes explicit probabilistic predictions about the ordering of propositions in human subject recall protocols in terms of the causal network and episodic memory network analysis of a given text. Samples from the TMF probability model were then used to generate synthetic protocol data using half of the human subject data. Statistics computed with respect to the remaining half of the human subject data and the synthesized protocol data were qualitatively similar in many respects. Relevant discrepancies between the human protocol data and synthesized protocol data were also identified.

The generic model of text comprehension and memory is based upon the idea that the reader forms a situation model as a by-product of the text comprehension process (Van Dijk & Kintsch, 1983). This situation model then plays a dominant role in guiding story recall, summarization, and question-answering processes. The components of a situation model can be modeled as knowledge digraphs (i.e., knowledge directed graphs or semantic networks). One example of such a knowledge digraph is the network generated by a story grammar (Stein & Glenn, 1979). Another closely related text knowledge digraph is a causal network that indicates how the propositions in a text are causally related to one another. It has been shown that statements with more causal connections are more likely to be recalled from memory (Trabasso, Secco, & Van den Broek, 1984; Trabasso & Van den Broek, 1985), included in a summary of a text (Van den Broek & Trabasso, 1986), and rated as more important (Van den Broek, 1988).

This paper discusses an application of the temporal Markov field (TMF) framework for constructing explana-

tory probabilistic situation models of recall protocol data. In this paper, only two text knowledge digraphs are considered. The first text knowledge digraph is a causal digraph that indicates which pairs of text propositions are causally related. The second text knowledge digraph is an episodic memory digraph that indicates the ordering of propositions in the original text.

This paper is organized in the following manner. First, the TMF framework is briefly described. Second, a text recall experiment involving human subjects is described. The data from the text recall experiment will then be used to illustrate an application of the TMF methodology for generating simulated text recall protocols from a given probabilistic text knowledge representation scheme.

TEMPORAL MARKOV FIELD FRAMEWORK

In this section, the details of the Markov field probability model are briefly reviewed. The specific formulas described in this section are discussed in detail and derived elsewhere (Golden, 1996a). An introduction to Markov random fields may be found in Golden (1996b).

Text Propositions

A particular text is modeled as an ordered sequence of text propositions. For example, consider the following story which is based upon the Aesop's fable titled "The Two Frogs":

The two frogs and the well. Two frogs dwelt in the same pool of water. The pool dried up in the summer's heat. The two frogs left the pool of water. The two frogs began their search for another pool of water. As the two frogs travelled along they reached a well. The well was very deep. The first frog said that the well would provide them with shel-

The author is grateful to Inah Choi, who assisted in writing the 14 texts used in this study, Dafna Yee, who assisted in collecting the human text recall data and the design of the human text recall data experiment, and to Frank Manganaro, who assisted in coding the text recall data. The author would also like to thank Neil Coady and Sean Ragan for their assistance in the development of the causal network analyses of the 12 texts used in this study, and Rebecca Horn and Mark Freeman for their ongoing encouragement, support, and feedback at various stages throughout this project. Also, the author thanks the School of Human Development at the University of Texas at Dallas for the necessary funds and support required to support this research. Correspondence should be addressed to R. M. Golden, Cognition and Neuroscience Program, School of Human Development (GR4.1), University of Texas at Dallas, Box 830688, Richardson, TX 75083-0688 (e-mail: golden@utdallas.edu).

ter and food. The first frog wanted to enter the well. The second frog replied, "Suppose the well's bottom is not a good home." The second frog said, "How will we get out of the deep well in that case?" (adapted from *The Book of Fables*, 1963)

A text proposition table explicitly identifies all text propositions within a given text. The text proposition table for "The Two Frogs" is shown in Table 1. Table 2 shows the text proposition table for "The Miser," which is based upon another fable written by Aesop (adapted from *The Book of Fables*, 1963). In the left-hand column of the text proposition table is the text proposition identification code that identifies the text proposition of interest. In the right-hand column is an illustrative instance of the text proposition concept.

In addition, a particular story text can be abstractly represented in terms of the elements of a text proposition table. The specific story text associated with the text proposition table in Table 1 is the *frog* story, which is abstractly represented by the following ordered sequence of text propositions:

$$\{f_1, f_2, f_3, f_5, f_6, f_7, f_8, f_9, f_{10}, f_{12}, f_{13}, f_{14}\}.$$

Similarly, human subject protocol data can also be represented within this framework as an ordered sequence of text propositions. Finally, notice that text propositions f_6 , f_{11} , and f_{12} are not present in the *miser* text. These text propositions are hidden text propositions and were used only for the purposes of coding human subject text recall protocol data. Hidden text propositions are indicated in the text proposition table as sentence fragments that are expressed in uppercase letters (see Table 2).

Working Memory Modeling Assumptions

Let y be a d -dimensional real vector whose j th element, y_j , indicates the degree to which text proposition j

Table 1
Frog Text Proposition Table

Text Proposition	Instance of Text Proposition (Sentence Fragment)
f_1	TITLE: The two frogs and the well.
f_2	Two frogs dwelt in the same pool of water.
f_3	The pool dried up in the summer's heat.
f_4	**THE TWO FROGS DECIDED TO FIND ANOTHER HOME**
f_5	The two frogs left the pool of water.
f_6	The two frogs began their search for another pool of water.
f_7	As the two frogs travelled along they reached a well.
f_8	The well was very deep.
f_9	The first frog said that the well would provide them with shelter and food.
f_{10}	The first frog wanted to enter the well.
f_{11}	**THE SECOND FROG WANTED TO THINK ABOUT THE WELL AS A NEW HOME**
f_{12}	The second frog replied, "Suppose the well's bottom is not a good home."
f_{13}	The second frog said, "How will we get out of the deep well in that case?"
f_{14}	The end.

Table 2
Miser Text Proposition Table

Text Proposition	Instance of Text Proposition (Sentence Fragment)
f_1	TITLE: The miser, the neighbor, and the gold.
f_2	A miser bought a lump of gold using all of his money.
f_3	The miser buried the gold in the ground.
f_4	The miser looked at the buried gold each day.
f_5	One of the miser's servants discovered the buried gold.
f_6	**THE SERVANT WANTED THE MISER'S GOLD**
f_7	The servant stole the gold.
f_8	The miser, on his next visit, found the hole empty.
f_9	The miser was very upset.
f_{10}	The miser pulled his hair.
f_{11}	**THE NEIGHBOR WANTED TO MAKE THE MISER FEEL BETTER**
f_{12}	**THE NEIGHBOR WANTED THE MISER TO REALIZE A STONE WAS AS USELESS AS THE GOLD**
f_{13}	A neighbor told the miser not to be upset.
f_{14}	The neighbor said, "Go and take a stone, and bury it in the hole."
f_{15}	The neighbor said, "And imagine that the gold is still lying there."
f_{16}	The neighbor said, "The stone will be as useful to you as the gold."
f_{17}	The neighbor said, "When you had the gold, you never used it."
f_{18}	The end.

is active in working memory. The vector y is called the *working memory buffer*.

Let the i th proposition in the text proposition table, f_i , be represented as a d -dimensional vector with a one in position i and zeros in all remaining $d-1$ positions. Let the notation $x(t) = f_i$ indicate that the t th proposition recalled by a human subject was text proposition f_i . Then, the working memory buffer y is updated according to the formula: $y(t+1) = x(t) + \mu y(t)$, where μ is some real number between 0 and 1 inclusive. The parameter μ is called the *working memory decay rate*. Notice that in the case where $\mu = 0$, information decays instantaneously in the working memory buffer. In practice, a small finite set of possible values for the working memory decay rate parameter is chosen, and the value of the working memory decay rate parameter that best fits the recall data is selected. In the analyses reported here, the set of possible values for μ were 0, 1/8, 2/8, 3/8, 4/8, 5/8, 6/8, and 7/8.

Knowledge Digraph Modeling Assumptions

The associative links identified by a text knowledge analysis are used to specify the free parameters of a special d -dimensional matrix, which is called the W matrix. The W matrix may be graphically represented as a weighted digraph (directed graph). The ij th element of W indicates the degree to which the activation of the j th text proposition, y_j , in the working memory buffer, $y = [y_1, \dots, y_d]$, influences the probability that the i th text proposition will be recalled.

This paper considers only two qualitative types of associative links in the W matrix. The first type of associative link is an episodic memory link, which is identi-

fied by examining the original order of presentation of text propositions in the texts presented to the human subjects. If text proposition i follows text proposition j , then an episodic memory link from proposition j to proposition i is formed. The second type of associative link is a causal link, which is identified using a causal network analysis based upon Mackie's (1980) notion of counterfactual causality. An overview of this causal network analysis may be found in the reviews by Van den Broek (1990a, 1990b). The details of the specific analysis used here are reviewed in a technical report by Golden (1996a). It is sometimes convenient to view the weighted digraph $\mathbf{W}$ as the sum of two weighted digraphs that correspond to the episodic memory link and causal link contributions.

The link strengths of both episodic memory and causal links in the weighted digraph $\mathbf{W}$ are free parameters of the model, which are estimated for each text using quasi-maximum likelihood estimation (see Golden, 1996b, for a discussion of quasi-maximum likelihood estimation). The remaining elements of the $\mathbf{W}$ matrix, which were neither episodic memory links nor causal links, were not free parameters and were constrained to have the value of zero. Example knowledge digraph analyses for the *frog* and *miser* texts are shown in Figures 1 and 2, respectively.

Probabilistic Modeling Assumptions

Let d be the number of text propositions in the text proposition table. As previously noted, the j th text proposition in the text proposition table is represented as a d -dimensional vector, where the j th element is equal to

Figure 1. The causal and episodic knowledge schemata associated with the *frog* text proposition table in Table 1. The number in each node identifies a particular text proposition in Table 1. A dashed arrow from node j to node i refers to the episodic relationship indicating that text proposition i immediately followed text proposition j in the text originally presented to the subject. A solid arrow from node j to node i refers to the causal relationship indicating that text proposition j caused the occurrence of text proposition i according to a counterfactual causality criterion. Causal and episodic associative link strengths whose values were estimated to be equal to zero are not shown in this figure.

Figure 2. The causal and episodic knowledge schemata associated with the *miser* text proposition table in Table 2. The number in each node identifies a particular text proposition in Table 2. A dashed arrow from node j to node i refers to the episodic relationship indicating that text proposition i immediately followed text proposition j in the text originally presented to the subject. A solid arrow from node j to node i refers to the causal relationship indicating that text proposition j caused the occurrence of text proposition i according to a counterfactual causality criterion. Causal and episodic associative link strengths whose values were estimated to be equal to zero are not shown in this figure.

one and all other $d-1$ elements are equal to zero. Assume that the human subject has recalled M text propositions $\mathbf{x}(1), \dots, \mathbf{x}(M)$ from memory, where $\mathbf{x}(j)$ is the text proposition that was located in the j th position of the recall protocol. For $t = 1 \dots M$, let

$$\mathbf{y}(t) = [y_1(t), \dots, y_d(t)] = \mathbf{x}(t-1) + \mu \mathbf{y}(t-1)$$

define the contents of the working memory buffer of the human subject after t propositions have been recalled. Assume $\mathbf{x}(0) = \mathbf{y}(0)$ is a d -dimensional vector of zeros. The real parameter μ ($0 \leq \mu < 1$) is the working memory decay rate.

The probability mass assigned to a given recall protocol for a given text with digraph $\mathbf{W}$ is defined by the formula

$$p(\mathbf{x}(1), \dots, \mathbf{x}(M) | M) = (1/Z) \exp \left[\sum_{t=1}^M \mathbf{x}(t)^T \mathbf{W} \mathbf{y}(t) \right], \quad (1)$$

where Z is some positive real-valued function of $\mathbf{W}$, which is not functionally dependent upon $\mathbf{x}(1), \dots, \mathbf{x}(M)$. For a given $\mathbf{W}$, Z is called the *normalization constant* and is chosen so that $p(\mathbf{x}(1), \dots, \mathbf{x}(M) | M)$ is a probability mass function.

The joint probability mass function in Formula 1 can now be used to calculate (see Golden, 1996a, for specific details) the local conditional probability the human subject will recall the i th text proposition in the text proposition table given knowledge of all other propositions recalled by the human subject (i.e., knowledge of propo-

sitions $1, 2, \dots, i-1, i+1, \dots, M$). In particular, let the local potential function $V_{s,i}$ be defined, such that

$$V_{s,i} = \mathbf{f}_i^T \left[\mathbf{W}\mathbf{y}(s) + \sum_{t=s+1}^M \left[\mathbf{W}\mu^{t-1-s} \right]^T \mathbf{x}(t) \right]. \quad (2)$$

Then, the conditional probability that the s th proposition in a recall protocol is the i th text proposition recalled given knowledge of all other propositions in the recall protocol can be explicitly computed from Formula 1 and shown (see Golden, 1996a, for details) to be expressed by the formula

$$p_{s,i} = \frac{\exp[V_{s,i}]}{\sum_{k=1}^d \exp[V_{s,k}]}. \quad (3)$$

TMF SAMPLING ALGORITHM EVALUATION STUDY

In this section, a specific application of the general TMF text schema framework to the problem of synthesizing human-like text recall protocols will be considered. Example applications of the TMF framework to text recall data analysis have already been described (Golden, 1994, 1995; Golden, Golden, Strickland, & Choi, 1993). The TMF sampling algorithm evaluation study involved two stages. The first stage (Experiment 1) involved the collection and analysis of human recall data. The second stage (Simulation Study 1) used half of the human recall data (parameter estimation data set) collected in the first stage to estimate the parameters of the TMF probability model and used the remaining human recall data (model evaluation data set) to evaluate both the TMF sampling algorithm and the text knowledge digraph analyses.

Experiment 1

In this experiment, 24 human subjects recalled 12 short paragraph-length texts from memory. The 24 subjects were then randomly divided into two groups of 12 subjects. Experiment 1 had three specific goals. First, one of the two groups of human subjects (the parameter estimation group) was used to estimate the parameters of the TMF sampling algorithm, and the other group of human subjects (the model evaluation group) was used to evaluate the sampling algorithm's performance. Thus, Experiment 1 provided a data set for evaluating the TMF sampling algorithm. Second, Experiment 1 was designed to replicate the classical finding that statements with more causal connections are more likely to be included in a recall protocol (Fletcher & Bloom, 1988; Graesser, 1978; Graesser, Robertson, Lovelace, & Swinehart, 1980; Trabasso et al., 1984; Trabasso & Van den Broek, 1985). And, third, Experiment 1 was designed to investigate the "stability" of text recall data for relatively small sample sizes. That is, to what extent are the qualitative and quantitative properties of human text recall data preserved across samples of small sample size?

Method

Materials. Fourteen short texts were used. The texts are referenced by the following 14 words: *bear, cuckoo, dragon, frog, hare, hunter, lion, merchant, miser, monkey, old woman, peasant, philosopher, and traveller*. The texts *bear* and *philosopher* were filler texts, whereas the remaining 12 texts were experimental texts. The texts were specially written adaptations of a set of fables and fairy tales (Anonymous, 1963; Walker, 1988). Specially designed "titles" for the 14 texts were also generated. The titles simply listed the main actors of a given text in order to avoid providing the subject with useful retrieval cues based upon the story's titles.

The adaptations were designed so that each sentence (statement) was identified with a unique complex text proposition. In addition, an attempt was made to keep the average number of words per sentence roughly constant while maintaining the simplest possible sentence structure. Sentences in the texts tended to follow a standard subject-verb-object format. In addition, the use of anaphoric reference in the texts was also minimized. Thus, the texts were designed so that there were few syntactic cues in the text for helping the reader recall one sentence given another sentence. Suppose that a sentence in the text had described an agent who is thinking about an event or who is verbally expressing an event. That sentence in the text was then rewritten to clearly identify the agent and the event using a standard subject-verb-object format. Representative examples of 2 of the 14 texts can be found in Tables 1 and 2.

Design. Twelve distinct sequences of all 14 texts were generated consisting of 12 experimental texts and 2 filler texts. The first text in all 12 sequences was the *bear* text. The last text in all 12 sequences was the *philosopher* text. The *bear* and *philosopher* texts were filler texts. The 12 text sequences were constructed using a balanced Latin square design so that, across all 12 sequences, (1) each experimental text was presented in all 12 possible positions in a sequence and (2) each experimental text preceded and followed each of the 11 other experimental texts.

Subjects. Twenty-four upper division students from the University of Texas at Dallas participated in the experiment in order to satisfy a course requirement. Pairs of subjects were randomly assigned to 12 groups. One member of each of the 12 subject pairs was randomly assigned to the *parameter estimation* group. The other member of that subject pair was assigned to the *model evaluation* group. Each of the 12 groups of subjects was randomly assigned 1 of the 12 text sequences. Each subject was individually tested by an experimenter.

Procedure. The experiment was divided into three parts. In the first part, each subject read each text in his or her assigned text sequence, one sentence at a time, using a self-paced procedure that involved clicking the mouse button on an arrow "button" on the computer screen. Each sentence was displayed in the center of the screen of a Macintosh SE30 computer. The computer program used in this experiment was written using the Hypercard programming language. In particular, the title of the text was first presented, then each sentence of the text was presented, and, finally, the phrase "The end" was presented by itself on the computer screen. The subjects were instructed to read the stories because they would be asked to answer questions about the stories at a later time. The subjects were not told they would be asked to recall the stories from memory.

During the second part of the experiment, the subjects participated in an intervening task, which was a pilot study for another experiment. During this task, the subjects read out loud several sets of unrelated sentences. This phase of the experiment lasted approximately 5–15 min.

During the third part of the experiment, the experimenter recorded all verbal responses of the subjects using a cassette tape recorder. The first sentence of each text (not the text title) was used as a retrieval cue. This sentence was displayed on the computer screen. While the sentence retrieval cue remained on the computer screen, the subjects were asked to verbally recall the original text from memory. The subjects were instructed to say "finished" or "done"

when they had completed their recall of the story. The sequence of texts presented to the subject during the retrieval phase of the experiment was exactly the same sequence presented to the subject during the reading/encoding phase of the experiment.

Results and Discussion

Human recall data coding. The recall protocol data were coded by one coder according to the following procedure. First, each subject's verbal protocol was transcribed. Second, the verbal protocol was parsed into phrases corresponding to the clauses identified by the text proposition table. Third, the verbal protocol was represented by a sequence of numbers indicating both the content and the ordering of the statements in the verbal protocol. For example, the first subject in the model evaluation group recalled the *miser* text as follows:

Okay this is where a miser got a big hole and he would put all his gold in and cover it up and he would go back and look at it all the time just to remind himself that he had this gold. Well evidently he told his neighbor and his neighbor one time went and got all the gold and then whenever he went back one day the gold wasn't there. So he was just devastated. He told his neighbor and the neighbor said why don't you put a stone in there and because you were never using the gold anyway and you can just go back and look at that stone and it would be the same.

This protocol was coded using the proposition table in Table 2 as the sequence of text proposition identification numbers

$f_3, f_4, f_8, f_9, f_{14}, f_{17}, f_{15}$.

Thus, for each of the 12 texts, 24 rows of text proposition identification numbers were generated corresponding to the recall protocols of the 24 subjects who recalled each of the 12 texts. The raw recall protocol data generated from the 12 subjects in the model evaluation group for the *frog* and *miser* texts are shown in Tables 3 and 4, respectively. Note that the human recall protocols were coded using both hidden and nonhidden text propositions.

Statement recall probability analyses across texts.

A statement's recall probability per protocol was defined as the reciprocal of the number of distinct statements in the protocol if the statement was recalled and zero otherwise. The statement recall probability data of the model evaluation group were analyzed using an analysis of variance (ANOVA) model. In addition, a causal network analysis of each of the 12 texts was done using the methodology of Van den Broek (1990a, 1990b). The details of the specific causal network analysis that was used may be found in the technical report by Golden (1996a).

The first factor (causal connectivity) had three levels that indicated whether a given statement participated in a small number of causal connections (i.e., zero or one causal connection), a moderate number of causal connections (i.e., two or three causal connections), or many causal connections (i.e., four or more causal connections). It should be noted that the distribution of causal connections is highly irregular, and, thus, the ANOVA model used to analyze the data is unbalanced. The sec-

Table 3
Human Recall (Model Evaluation Group)
Protocol Data for *Frog* Text

Subject I.D.	Text Proposition Sequence
Subject 1	f_3, f_4, f_7, f_{12}
Subject 2	$f_3, f_6, f_7, f_9, f_{12}, f_{13}$
Subject 3	$f_3, f_7, f_8, f_{10}, f_9, f_{12}, f_{13}$
Subject 4	f_6, f_7, f_{10}, f_{12}
Subject 5	f_7, f_8, f_{12}, f_{13}
Subject 6	f_3, f_4
Subject 7	f_4, f_3
Subject 8	f_3, f_7, f_9, f_{11}
Subject 9	f_{10}
Subject 10	f_3
Subject 11	$f_3, f_6, f_7, f_{10}, f_{12}$
Subject 12	$f_3, f_6, f_7, f_9, f_{10}, f_{12}, f_{13}$

Table 4
Human Recall (Model Evaluation Group)
Protocol Data for *Miser* Text

Subject I.D.	Text Proposition Sequence
Subject 1	$f_3, f_4, f_8, f_9, f_{14}, f_{17}, f_{15}$
Subject 2	$f_3, f_4, f_7, f_9, f_{14}, f_{15}, f_{16}$
Subject 3	$f_3, f_4, f_5, f_7, f_9, f_{14}, f_{15}, f_{16}$
Subject 4	$f_3, f_5, f_7, f_{14}, f_{15}, f_{16}$
Subject 5	f_7, f_4, f_{14}
Subject 6	$f_3, f_4, f_7, f_8, f_9, f_{14}, f_{17}$
Subject 7	$f_3, f_7, f_{14}, f_{16}, f_{17}$
Subject 8	$f_3, f_4, f_5, f_8, f_9, f_{14}, f_{17}$
Subject 9	$f_3, f_8, f_{12}, f_{14}, f_{16}, f_{17}, f_{15}, f_{17}$
Subject 10	$f_3, f_5, f_7, f_{14}, f_{16}$
Subject 11	$f_3, f_4, f_7, f_9, f_{14}, f_{16}, f_{17}$
Subject 12	$f_3, f_4, f_5, f_7, f_8, f_9, f_{14}, f_{15}, f_{17}, f_{16}$

ond factor (text) indicated the story identity and, thus, had 12 levels. The third factor (statement item) was the identity of a particular statement that had originally appeared in the text and, thus, could potentially be recalled by the subject. The item factor was nested within the crossing of the text and causal connectivity factors. The item factor was also crossed with the subjects factor. The fourth factor was the subjects factor.

The analysis was done in three different ways since there were three random factors (subjects, texts, and items). First, the text and item factors were ignored, and the subjects factor was treated as a random factor. The F statistic computed for this analysis will be referred to as F_s . Second, the subjects and item factors were ignored, and the text factor was treated as a random factor. The F statistic computed for this analysis will be referred to as F_t . Third, the subject and text factors were ignored, and the items factor was treated as a random factor. The F statistic computed for this analysis will be referred to as F_i . The statistic $F_{\min}$ is computed by taking the minimum of F_s , F_t , and F_i .

The data analysis of the human subject model evaluation group replicated the well-known finding that statements with more causal connections are more likely to be recalled. In particular, as the number of causal connections to a statement increased, that statement's recall probability tended to increase as well [$F_s(2,2) = 31.52$,

$p < .0001$; $F_1(38,1557) = 4.17$, $p < .0001$; $F_{\min}(2,22) = 12.09$, $p < .0005$].

Statement recall probabilities for individual texts.

In this section, an additional detailed analysis of statement recall probabilities for individual texts was done. The degree to which the two groups of human subjects provided consistent response patterns was first evaluated. Statement recall probability for the first group of human subjects (parameter estimation group) were positively correlated with statement recall probability for the second group of human subjects (model evaluation group) for 9 of the 12 texts. These 9 texts were *cuckoo* [$r(10) = .682$, $p < .05$], *dragon* [$r(12) = .94$, $p < .05$], *hunter* [$r(11) = .81$, $p < .05$], *lion* [$r(9) = .64$, $p < .05$], *merchant* [$r(8) = .73$, $p < .05$], *miser* [$r(10) = .93$, $p < .05$], *monkey* [$r(8) = .66$, $p < .05$], *old woman* [$r(8) = .81$, $p < .05$], and *peasant* [$r(12) = .84$, $p < .05$]. Statement recall probability for the first group of human subjects was not correlated with the second group of human subjects for the remaining three texts: *frog* [$r(7) = .35$, $p = .32$], *hare* [$r(7) = .48$, $p = .15$], and *traveller* [$r(7) = .41$, $p = .24$]. Despite the fact that each group of subjects consisted of only 12 human subjects, there was fairly good agreement between patterns of responses for statement recall probabilities for specific texts.

Next, an analysis was done to evaluate the extent to which statement recall probability for a given statement was correlated with the number of causal connections associated with that statement for the human subjects in the model evaluation group. For the model evaluation group of human subjects, statement recall probability was positively correlated with causal connectivity for 7 of the 12 texts: *dragon* [$r(12) = .61$, $p < .05$], *hunter* [$r(11) = .67$, $p < .05$], *lion* [$r(9) = .55$, $p < .05$], *merchant* [$r(8) = .81$, $p < .05$], *miser* [$r(10) = .61$, $p < .05$], *old woman* [$r(8) = .87$, $p < .05$], and *peasant* [$r(12) = .58$, $p < .05$]. Statement recall probability was not correlated with causal connectivity for the other 5 texts: *cuckoo* [$r(10) = .26$, $p = .39$], *frog* [$r(7) = .16$, $p = .66$], *hare* [$r(7) = .14$, $p = .71$], *monkey* [$r(8) = .41$, $p = .20$], and *traveller* [$r(7) = .18$, $p = .63$].

Simulation Study 1

In this computer simulation study, the data collected in Experiment 1 from the parameter estimation group were used to estimate the parameters of the TMF sampling algorithm. The data from the model evaluation group of human subjects were used to evaluate the TMF sampling algorithm's performance. There were four major goals for this simulation study.

First, Simulation Study 1 was designed to evaluate the TMF synthesized recall protocols in order to see if statements with more causal connections were more likely to be included in synthesized TMF recall protocols. Thus, a qualitative comparison between the TMF synthesized recall data and human recall data could be made. Second, because the parameters of the TMF model were estimated from one group of human subjects and the TMF

sampling algorithm was tested on a second group of subjects, novel quantitative predictions of the TMF sampling algorithm could be made.

A third major goal of this simulation study was to evaluate the "stability" of text recall data synthesized by the TMF model. In particular, the issue of how the TMF model's predictions varied as a function of the number of iterations of the algorithm was systematically investigated. According to the main convergence theorem (see Geman & Geman, 1984; Golden, 1996a), after a sufficiently large number of iterations of the Gibbs Sampler algorithm, the TMF synthesized recall protocols will be representative of true random samples from the TMF probability distribution. To investigate the rate of convergence (i.e., the validity of the synthesized recall protocols), synthesized recall protocols generated from 10, 100, and 1,000 iterations of the sampling algorithm were examined in this simulation study.

Finally, the fourth major goal of the simulation study was to closely examine a few of the temporal ordering predictions of the TMF model. Unlike other models of text recall in the literature (e.g., Fletcher & Bloom, 1988; Graesser et al., 1980; Kintsch & Van Dijk, 1978), the TMF model can make explicit qualitative and quantitative predictions regarding the ordered sequence of statements recalled by human subjects.

Method

Texts and knowledge digraphs. The texts that were used in Experiment 1 and the knowledge schemata and recall data generated in Experiment 1 were used in this simulation study.

Parameter estimation procedure. Approximately 1.1 h of CPU time on a Sparc 20 are required in order to estimate the episodic memory and causal digraph parameters for the eight *miser* models for the eight possible values of the working memory decay rate. In this prototype version of the program that is implemented in the MATLAB programming language, no attempt was made to "optimize" either the algorithm (which was a generic gradient descent type algorithm) or the software code. Details of the specific gradient descent optimization procedure may be found in Golden (1996a).

Procedure for generation of synthesized recall protocols. Since there were 12 subjects in the human training data group and 12 subjects in the human testing data group, 12 recall protocols were synthesized for each of the 12 texts. The procedure for generating the synthesized recall protocols involves using Equation 3 in conjunction with the Geman and Geman (1984) Gibbs sampler algorithm (see Golden, 1996b, for a review). The procedure is summarized as follows.

Step 1. Sample from the empirical distribution of recall protocols in order to obtain the number of items to be recalled by the simulated human subject. Let this number of items be denoted as M .

Step 2. Note that each text proposition is a d -dimensional vector with a 1 in exactly one position (indicating the text proposition's identity) and zeros in the remaining $d-1$ positions. Generate a recall protocol consisting of M randomly chosen and ordered text propositions given the constraint that the 1st and M th text propositions are chosen to be the retrieval cue given to the human subjects and the last text proposition of the story ("The end"), respectively.

Step 3. Pick a random integer from the set $\{2, 3, \dots, M-1\}$. Denote this random integer as s .

Step 4. Use Equation 3 to compute the local conditional probability of the s th item in the recall protocol being equal to the i th

text proposition. Then, let the s th item in the recall protocol be equal to a sample from that local conditional probability mass function.

Step 5. Go to Step 3 until some convergence criteria are satisfied.

The Geman and Geman (1984) Gibbs sampler algorithm for the TMF probability model was used (see Appendix of Golden, 1996a, for the details of the derivation) to sample from the TMF probability distribution. The algorithm is a natural generalization of a Boltzmann machine algorithm (Ackley, Hinton, & Sejnowski, 1985) running at "thermal equilibrium" (i.e., temperature equal to one) where the units are d -valued (where d is the number of text propositions) as opposed to binary-valued units. In this prototype version of the program that is implemented in the MATLAB programming language, no attempt was made to optimize the software code. Approximately 2.8 h of cpu time on a Sparc 20 are required in order to generate 24 synthetic recall protocols from the *miser* TMF probability model.

Results and Discussion

Statement recall probability analyses across texts.

As in Experiment 1, a statement's recall probability per protocol was defined as the reciprocal of the number of distinct statements in the synthesized protocol if the statement was recalled and zero otherwise. Figure 3 provides a comparison of the performance of the TMF sampling algorithm after 1,000 iterations with the human subject data. For statement recall probabilities across texts, the TMF model provides a reasonable fit to the human subject parameter estimation data. In addition, the TMF model makes surprisingly effective qualitative and quantitative predictions on the human subject model evaluation data set. Note that both human subject data sets and the TMF-generated data exhibit the important qualitative feature that, as causal connectivity increases, statement probability of recall tends to increase. Figure 3 also provides insights into the intrinsic variability in the human subject protocol data; the data from two independent groups of human subjects are plotted.

Figure 3. Human and model statement recall probability as a function of causal connectivity. Statistics are computed across all 12 texts. The human training (parameter estimation) data and human test (model evaluation) data are both qualitatively and quantitatively similar to the TMF model's synthesized protocol data. The human and model protocol data are consistent with the hypothesis that statements with more causal connections are more likely to be included in a recall protocol.

Figure 4 shows statistics computed from the TMF model's synthesized recall protocols after 10, 100, and 1,000 iterations of the sampling algorithm. Notice that the statement recall probability statistics are relatively stable between 100 and 1,000 iterations of the TMF sampling algorithm. After only 10 iterations, however, the statement recall probability statistics are not completely stable. In particular, statements with exceptionally low statement recall probabilities tended to have their recall probabilities overestimated by the sampling algorithm after only 10 iterations.

The statement recall probability data generated by the TMF sampling algorithm were analyzed using an ANOVA model where the first factor (iterations) was a between-groups factor whose three levels indicated if the data were generated from a simulation involving 10, 100, or 1,000 TMF algorithm iterations. The second factor (causal connectivity) had three levels that indicated whether a given statement participated in a small number of causal connections (i.e., zero or one causal connection), a moderate number of causal connections (i.e., two or three causal connections), or many causal connections (i.e., four or more causal connections). As in Experiment 1, the ANOVA model used to analyze the design was unbalanced because no attempt was made to control the causal connectivity structure of the texts. A third factor (text) indicated the story identity and, thus, had 12 levels. The fourth factor (statement item) indicated the identity of a statement that had originally been presented in a particular text and could potentially be recalled. The item factor was nested within the crossing of the text factor and the causal connectivity factor. The fifth factor (computer simulated subjects) was nested within the iterations factor. The subjects factor was crossed with the item factor. The $F_{\min}$ analysis of Experiment 1 was used to analyze the data in the computer simulations as well.

As the number of causal connections to a statement increased, that statement's recall probability tended to increase as well [$F_s(2,66) = 885.08, p < .0001$; $F_t(2,22) = 24.81, p < .0001$; $F_{\min}(2,36) = 13.25, p < .0001$]. In addition, the statement recall probabilities for the 10 iteration condition ($M = .06640$) were different from the statement recall probabilities in the 100 iteration ($M = .07303$) and the 1,000 iteration ($M = .07307$) conditions [$F_t(2,22) = 9.11, p < .005$; $F_s(2,33) = 8.97, p < .005$; $F_{\min}(2,72) = 3.19, p < .05$]. There was no interaction of the iteration and causal connectivity conditions [$F_s(4,66) = 2.78, p < .05$; $F_t(4,72) = 0.92, p = .46$; $F_{\min}(4,44) = 0.76, p = .56$].

Quantitative and qualitative predictions for specific texts. This section examines the predictive performance of the model with respect to statement recall probabilities for specific texts. In particular, statement recall probabilities for the synthesized TMF model recall data set were positively correlated with statement recall probabilities for the human subjects (model evaluation group) for 6 of the 12 texts: *dragon* [$r(12) = .59, p < .05$], *hunter* [$r(11) = .54, p < .05$], *lion* [$r(9) = .70, p < .05$], *miser* [$r(10) = .87, p < .05$], *old woman* [$r(8) = .66, p < .05$],

Figure 4. Statement recall probability as a function of causal connectivity and TMF algorithm iterations. The statistics computed from the TMF synthesized recall protocols are similar for 100 and 1,000 iterations of the algorithm, which is not inconsistent with the hypothesis that the TMF algorithm has converged.

and *peasant* [$r(12) = .65, p < .05$]. Statement recall probabilities for the synthesized TMF model recall data set were not correlated with the human subject (model evaluation group) data for the other 6 texts: *cuckoo* [$r(10) = .25, p = .42$], *frog* [$r(7) = .022, p = .96$], *hare* [$r(7) = .35, p = .33$], *merchant* [$r(8) = .54, p = .07$], *monkey* [$r(8) = .35, p = .29$], and *traveller* [$r(7) = .56, p = .07$].

Figure 5 shows how statement recall probabilities for the TMF model and both groups of human subjects vary as a function of causal connectivity for the *frog* text. The *frog* text is distinguished as having an especially low correlation between the model's recall and the recall of human subjects in the model evaluation group. Examination of Figure 5 still reveals a fairly good qualitative match between the model and human subjects' performance. Note that the statements on the x-axis of Figure 5 are ordered from left to right according to their degree of causal connectivity. Statements with less causal connectivity are located toward the left and statements with greater causal connectivity are located toward the right.

Figure 6 shows how statement recall probabilities for the TMF model and both groups of human subjects vary as a function of causal connectivity for the *miser* text. The *miser* text is distinguished as having an especially high correlation between the model's recall and the recall of human subjects in the model evaluation group. Examination of Figure 6 reveals an excellent qualitative match between the model and human subjects' performance. In addition, the estimated recall probabilities agree quantitatively with the human recall probabilities from the model evaluation group.

Comparison of synthesized and human recall protocol data. Although the probability that a statement is included in a protocol is a commonly used dependent measure in the literature, this dependent measure deliberately ignores the sequential or temporal organization of human protocol data. For example, a theoretical model

that is limited to predicting the likelihood that a statement will be included in a protocol would be unable to distinguish the difference between two protocols that contain exactly the same text propositions arranged in different orders. The TMF model, however, makes explicit predictions regarding the order in which statements will be arranged in a protocol.

Tables 5 and 6 show the synthesized TMF generated recall protocol data for the *frog* and *miser* texts, respectively. Statement recall probabilities were relatively weakly correlated for the *frog* text and relatively strongly correlated for the *miser* text. Thus, the *frog* text is representative of a relatively poor fit to the human data, whereas the *miser* text is representative of a relatively good fit. The reader is encouraged to compare the synthesized TMF protocol data in Tables 5 and 6 with the human model evaluation group protocol data in Tables 3 and 4.

For example, note that the sixth synthesized protocol for the *miser* text in Table 6 is the proposition sequence $f_3, f_7, f_{14}, f_{16}, f_{17}$, which is exactly the same as the seventh protocol for the *miser* text in Table 4. Moreover, this protocol was not present in the human subject parameter estimation data set and, thus, was not directly learned by the model. As previously noted, this text proposition sequence can be translated into English by examining the text proposition table for the *miser* text. Thus, $f_3, f_7, f_{14}, f_{16}, f_{17}$ may be represented as the sequence of English statements:

The miser buried the gold in the ground. The servant stole the gold. The neighbor said, "Go and take a stone, and bury it in the hole." The neighbor said, "The stone will be as useful to you as the gold." The neighbor said, "When you had the gold, you never used it."

Figure 5. Human and model statement recall probability as a function of causal connectivity for the *frog* text. The human training (parameter estimation) data and human test (model evaluation) data are both qualitatively and quantitatively similar to the TMF model's synthesized protocol data. Note that the model (synthesized protocol) protocol data and human train (parameter estimation) protocol data are consistent with the hypothesis that statements with more causal connections are more likely to be included in a recall protocol. However, this trend is not strongly evident in the human test (model evaluation) data.

Figure 6. Human and model statement recall probability as a function of causal connectivity for the *miser* text. The human training (parameter estimation) data and human test (model evaluation) data are both qualitatively and quantitatively similar to the TMF model's synthesized protocol data. Note, in particular, that both the human and the model protocol data are consistent with the hypothesis that statements with more causal connections are more likely to be included in a recall protocol.

One important qualitative difference between synthesized recall protocol data and human recall data is proposition doubling. An example of proposition doubling is shown in the second synthesized protocol for the *frog* text. Notice that this protocol $f_3, f_6, f_9, f_9, f_{12}, f_{13}$ is quite similar to the human subject data in Table 5, with the exception that text proposition f_9 is repeated twice by the model. Although one could assume the existence of an additional output editing process that ignores text proposition that has already just been recalled, such an assumption still does not explain this peculiar characteristic of the model's performance unless the details of such an editing process are worked out.

In addition, the model exhibits other phenomena—besides proposition doubling—that occur much more frequently in the synthesized recall data relative to the human recall data. For example, consider protocol 12 of the synthesized *miser* text. The first four text propositions synthesized by the model are f_3, f_5, f_3, f_5 . This phenomenon will be called the *subsequence repetition* phenomenon.

Another type of qualitative discrepancy between the human and model recall data can arise due to the absence of associative links. For example, in Protocol 12 of the human subject data in Table 3, Text Proposition 12 is recalled following the recall of Text Proposition 10. This transition is not observable in the synthesized recall data in Table 5. The reason for the absence of this transition is most likely due to the fact that there is not an episodic or causal link from Text Proposition 10 to Text Proposition 12 in the *frog* knowledge schema digraphs (see Figure 1).

Of course, the absence of such links does not imply that the TMF probability model could never generate Text Proposition 12 followed by Text Proposition 10. For ex-

ample, in the fourth synthesized protocol of Table 5, Text Proposition 10 is immediately followed by Text Proposition 3. Inspection of Figure 5 shows that there are no associative links from Text Proposition 3 to Text Proposition 10. Due to the complex interactions among the various associative link strengths and working memory assumptions in conjunction with the fact that the recall mechanism in the model is stochastic, it is sometimes difficult to “explain” observed transitions.

On the other hand, consider the transition from Text Proposition 10 to Text Proposition 6 in synthesized protocol four of Table 5. Just before Text Proposition 10 is Text Proposition 3, which does have a causal associative link to Text Proposition 6. In addition, Text Proposition 6 has a causal associative link to Text Proposition 12, which follows Text Proposition 6 in the protocol. Thus, because the TMF model uses information about all other items in the protocol to recall a specific item, the location of Text Proposition 6 in the protocol is probably due to the interaction of the working memory decay rate assumption and the associative links from Text Proposition 6 to Text Propositions 3 and 12.

Other interesting discrepancies between the human and model recall data also exist that cannot be easily explained in terms of the knowledge structure digraphs. For example, in the human recall protocol data for the *frog* text (Table 3), Text Proposition 7 is preceded frequently by either a 3, 4, or 6. In the synthesized recall protocol

Table 5
TMF Synthesized Recall
Protocol Data (1,000 Iterations) for *Frog* Text

Subject I.D.	Text Proposition Sequence
Simulation 1	f_9, f_{12}, f_{13}
Simulation 2	$f_3, f_6, f_9, f_9, f_{12}, f_{13}$
Simulation 3	$f_{12}, f_{13}, f_8, f_9, f_{12}, f_{13}$
Simulation 4	$f_3, f_{10}, f_6, f_{12}, f_{13}, f_{12}, f_{13}$
Simulation 5	$f_3, f_6, f_9, f_9, f_{12}, f_{12}, f_{13}$
Simulation 6	f_6, f_9, f_{12}, f_{13}
Simulation 7	$f_3, f_6, f_9, f_{12}, f_{13}, f_{12}, f_{13}$
Simulation 8	$f_3, f_6, f_9, f_{12}, f_{13}, f_{12}, f_{13}$
Simulation 9	f_6, f_9, f_{12}, f_{13}
Simulation 10	$f_6, f_9, f_{12}, f_{13}, f_{13}, f_9, f_{12}, f_{12}, f_{13}$
Simulation 11	$f_3, f_6, f_9, f_{12}, f_9, f_{12}, f_{13}$
Simulation 12	$f_3, f_6, f_{12}, f_9, f_{12}, f_{13}$

Table 6
TMF Synthesized Recall
Protocol Data (1,000 Iterations) for *Miser* Text

Subject I.D.	Text Proposition Sequence
Simulation 1	$f_3, f_3, f_5, f_7, f_8, f_7, f_9, f_9, f_{14}, f_{16}, f_{17}$
Simulation 2	$f_3, f_7, f_9, f_{14}, f_{14}, f_{15}, f_{16}, f_{17}, f_{16}, f_{16}, f_{17}$
Simulation 3	$f_3, f_4, f_6, f_7, f_9, f_{14}, f_{14}, f_{16}, f_{17}$
Simulation 4	$f_3, f_7, f_8, f_9, f_{14}, f_{14}, f_{16}, f_{17}, f_{16}$
Simulation 5	$f_3, f_3, f_5, f_7, f_7, f_9, f_{14}, f_{14}, f_{16}$
Simulation 6	$f_3, f_7, f_{14}, f_{16}, f_{17}$
Simulation 7	$f_3, f_7, f_7, f_9, f_{14}, f_{16}, f_{17}$
Simulation 8	$f_3, f_5, f_7, f_{14}, f_{14}, f_{16}$
Simulation 9	$f_3, f_3, f_5, f_7, f_7, f_{14}, f_{16}, f_{14}, f_{15}, f_{16}, f_{17}$
Simulation 10	$f_3, f_5, f_7, f_8, f_9, f_{14}, f_{14}, f_{16}, f_{17}$
Simulation 11	$f_3, f_3, f_5, f_7, f_8, f_9, f_{14}, f_{14}, f_{15}, f_{16}, f_{17}$
Simulation 12	$f_3, f_5, f_3, f_5, f_7, f_8, f_{14}, f_{16}$

data for the *frog* text (Table 5), however, Text Proposition 7 never occurs in the protocol data. The reason for this specific pattern of results is currently unknown.

Finally, human subjects can recall items that were never presented in the original texts. For example, several human subjects (see Table 3) recalled Text Proposition 4, which was an important *frog* goal ("The two frogs decided to find another home"), but, of course, this text proposition was not recalled by the model since there were no episodic or causal relationships linked to this hidden text proposition. Improved theories of semantic relationships among both visible and hidden elements in a text will be helpful for addressing this modeling problem.

GENERAL SUMMARY AND DISCUSSION

This paper reviewed the TMF text schema modeling framework for analyzing text protocol data as a categorical time series and showed how the Gibbs sampler algorithm (Geman & Geman, 1984) could be used as a method for sampling from the high-dimensional TMF probability distribution.

To empirically investigate and illustrate the ability of the TMF sampling algorithm to computationally evaluate causal network analyses, recall protocol data was collected from a group of human subjects. The recall data supported the well-known finding that statements with more causal connections are more likely to be recalled. An analysis of the synthesized recall protocol data generated from the TMF sampling algorithm and the causal network analyses revealed similar support for this established finding. In addition, the TMF sampling algorithm and causal network analyses made reasonably good quantitative predictions about statement recall probabilities for several specific texts. Finally, the detailed temporal structure of human and synthesized recall protocols was compared in a final data analysis.

Some noticeable and important discrepancies (e.g., proposition doubling and subsequence repetition) between the synthesized recall protocols and the human recall data were observed and noted. Ideally, the observation of such discrepancies could provide important insights into how to improve the causal network analyses and insights into what additional text knowledge digraphs (besides causal network and episodic memory network digraphs) are required to build computationally effective models of temporal structural regularities in recall protocol data. It should be emphasized, again, that the TMF framework is neither a model of how recall protocols are generated nor a specific theory of text knowledge representation. Rather, the TMF framework is a sophisticated theoretical framework for empirically investigating and developing text knowledge digraph analyses using the sequential structure of protocol data.

REFERENCES

- ACKLEY, D. H., HINTON, G. E., & SEJNOWSKI, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science*, 9, 147-169.
- THE BOOK OF FABLES (1963). New York: Frederick Warne.
- FLETCHER, C. R., & BLOOM, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory & Language*, 27, 235-244.
- GEMAN, S., & GEMAN, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 6, 721-741.
- GOLDEN, R. M. (1994). Analysis of categorical time-series text recall data using a connectionist model. *Journal of Biological Systems*, 2, 283-305.
- GOLDEN, R. M. (1995). Making correct statistical inferences using a wrong probability model. *Journal of Mathematical Psychology*, 38, 3-20.
- GOLDEN, R. M. (1996a). *A connectionist temporal Markov field model of text recall* (Tech. Rep. 96-01). Richardson, TX: University of Texas at Dallas, School of Human Development.
- GOLDEN, R. M. (1996b). *Mathematical methods for neural network analysis and design*. Cambridge, MA: MIT Press.
- GOLDEN, R. M., GOLDEN, S. F., STRICKLAND, J., & CHOI, I. (1993). A psychometric PDP model of temporal structure in story recall. In *The Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 487-491). Hillsdale, NJ: Erlbaum.
- GRAESSER, A. C. (1978). How to catch a fish: The memory and representation of common procedures. *Discourse Processes*, 1, 72-89.
- GRAESSER, A. C., ROBERTSON, S. P., LOVELACE, E. R., & SWINEHART, D. M. (1980). Answers to why questions expose the organization of story plot and predict recall of actions. *Journal of Verbal Learning & Verbal Behavior*, 19, 110-119.
- KINTSCH, W., & VAN DIJK, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- MACKIE, J. L. (1980). *The cement of the universe: A study of causation*. Oxford: Oxford University Press, Clarendon Press.
- STEIN, N. L., & GLENN, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing: Vol. 2. Advances in discourse processes* (pp. 53-120). Norwood, NJ: Ablex.
- 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). Hillsdale, NJ: Erlbaum.
- TRABASSO, T., & VAN DEN BROEK, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory & Language*, 24, 612-630.
- VAN DEN BROEK, P. (1988). The effects of causal relations and hierarchical position on the importance of story statements. *Journal of Memory & Language*, 27, 1-22.
- VAN DEN BROEK, P. (1990a). 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. (1990b). Causal inferences and the comprehension of narrative texts. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation* (Vol. 25, pp. 175-196). Hillsdale, NJ: Erlbaum.
- VAN DEN BROEK, P., & TRABASSO, T. (1986). Causal networks versus goal hierarchies in summarizing text. *Discourse Processes*, 9, 1-15.
- VAN DIJK, T. A., & KINTSCH, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- WALKER, B. K. (1988). *A treasury of Turkish folktales for children*. Hamden, CT: Shoe String Press.