JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 18, 309–318 (1979)

Episodes as Chunks in Narrative Memory

JOHN B. BLACK

University of Illinois at Chicago Circle

AND

GORDON H. BOWER

Stanford University

Story statements cluster into episodes, so we expected the memory representations of statements to cluster into separate episode chunks in memory. Experiment 1 confirmed this chunking idea by showing that the recall of episode actions depends on the length of that episode, but not on the lengths of other episodes. Specifically, adding more actions to an episode increased recall of the original actions but didn't affect recall of other episodes. Further, subjects rated the original actions as more important than the added actions, thus suggesting that unimportant actions increased the recall of important actions in that episode. Another experiment showed that the more subordinate actions an episode contains, the more likely a statement summarizing that episode is to be recalled.

Stories have two characteristics that distinguish them from unrelated groups of sentences. First, the sentences in a story fit together meaningfully: they are connected by cohesive relations into a coherent whole. However, these interconnected statements do not form a uniform network; rather the statements tend to cluster into subgroups. Thus the second distinguishing characteristic of stories is that they have a substructure or constituent structure. The experiments reported in this paper investigate the psychological reality of these constituents as chunks in people's memory representation for stories.

The theories of story memory that have emphasized the constituent structure of stories are the "story grammar" theories of Rumelhart (1975, 1977), Meyer (1975), Thorndyke (1977), Mandler and Johnson (1977), and Stein and Glenn (in press). Although differing in details, these theories all first divide a story into a setting constituent and a plot constituent. They also all divide the plot into a series of episodes and they define an episode as a subgoal, the actions that attempt to obtain that subgoal, and the outcome of those actions. Hence to validate this class of theories we should look for a psychologically relevant division between the setting and the plot of a story, and for a subdivision of the plot into goal-oriented episodes.

There are also two other theories that contain goal-oriented episodes: namely, the "macro-structure" theory of van Dijk and Kintsch (van Dijk, 1977; Kintsch, 1977b; Kintsch & van Dijk, 1978) and the "schema" theory of Schank and Abelson (1977). Both of these theories have a microscopic level that represents the events in the story and a macroscopic level that describes the events at

We thank Gwyn Felt for her valuable assistance with the experiments. This research was supported by Research Grant MH-13950 to the second author from the National Institutes of Mental Health. Reprints may be obtained from Dr. Gordon H. Bower, Department of Psychology, Stanford University, Stanford, California 94305.

0022-5371/79/030309-10$02.00/0

a more global or general level. The macroscopic level of the van Dijk and Kintsch theory makes the same setting versus plot and plot episode constituent distinctions as the "story grammars." However, the Schank and Abelson macroscopic level only makes the plot episode distinction.

The story memory theories which do not have these constituents are Kintsch (1974), Schank (1975), Frederiksen (1972, 1975a, 1975b), and Frederiksen (1975c, 1977). These theories have only a microscopic level of description.

If it is true that the statements in a story become filed or catalogued in memory under different constituent categories, like separate "chunks" in memory, we should inquire into the implications of such an organization for memory behavior. One implication is that the material in a chunk should act somewhat like "all-or-none" units in recall. That is, all material in one chunk will tend to be recalled together and somewhat independently of recall of the material in other chunks. This means if we cue recall with an action or proposition from one constituent, the subject is likely to retrieve other propositions from the same constituent. Thus, for example, if we cue the learner with an action from one episode of a multiepisode story, he is most likely to recall the goal, outcome, and other actions in that episode rather than information occurring in another episode. We know of no direct evidence for this prediction, though it seems intuitively correct.

Taking a suggestion from memory theories (e.g., Johnson, 1970; Mandler, 1967), another implication of our chunking hypothesis is that recall of material is limited by the number of chunks in the material, and not so much by the amount of material in each chunk. Ordinarily, the recall probability of each proposition decreases as the amount of material increases because increasing material to-be-learned usually means more chunks—the so-called "list length effect." However, recall probability does not decline if the added material falls in the same number of other chunks (e.g., see Bower, 1969). We will call this the "chunk-independence" hypothesis.

Consider one illustration. Johnson (1970) showed that subjects' learning of one group within a chunked letter series say, the second group in the series (BX) (PLZ) (HTL)] decreased with the size of that group and the number of other groups, but was independent of the size of the other groups. As a second illustration, Tulving and Pearlstone (1966) showed that free recall of the instances of a taxonomic category decreased as more categories were added to the list, but increasing the total number of items had no effect if the number of categories remained constant.

The question motivating the current experiments was whether we could show comparable effects in story memory of varying the size of the distinct "chunks" or constituents of a narrative. The basic idea was to take stories with two clearly distinct constituents, and vary the number of text propositions stated in the two constituents of the story. The critical prediction from the chunk-independence hypothesis is that recall of the propositions in a given constituent will be independent of the number of propositions in the other constituent of the story. In other words, increasing the propositions in a story without increasing the number of higher-level constituents should not affect recall of a target set of propositions in the story.

Although recall is expected to be independent of other chunk sizes, it is not clear what to predict for recall as a function of the size of the target chunk. Prior data give no unequivocal prediction. The multiple items within a chunk can be conceived to facilitate recall of the chunk itself but to create competition and interference among the multiple candidates for learning and output. For example, Tulving and Pearlstone (1966) found that increasing the number of items in a category increased the probability that something from the category would be recalled, but decreased the probability that any given category item

would be recalled. Therefore, we had no firm expectation about the net effect of varying items per chunk upon recall rate within that chunk.

First, we conducted a pilot study to test the feasibility of the separation into setting and plot constituents. The idea was that the length of the setting constituent would affect the free recall percentage of the setting statements but not the plot statements, and that the number of plot statements would affect the free recall percentage of the plot statements but not the setting statements. Instead we found that the length of the setting had an effect on the recall of both setting and plot statements, while the length of the plot had no effect on recall of either (although there was also an interaction of the setting and plot lengths).

This pilot study cast doubt on the separation of setting and plot into separate constituents and convinced us to concentrate on the plot episode constituents. Hence, in Experiment 1 we varied the length of two goal-oriented plot episodes to see if these proposed constituents showed the recall independence property.

Experiment 1

Method

Subjects. The subjects were 32 Stanford undergraduates either fulfilling a service requirement for their Introductory Psychology class or receiving payment.

Materials. Each subject read four different stories. The goals of the different characters in the four stories were to learn to play tennis, to find a book for a college class, to make money investing, and to find a job. Each story involved two "attempt episodes," methods the character tried to achieve this goal. For example, in looking for the book, the first episode described looking in the library (unsuccessfully) and the second described looking in the college bookstore (successfully). Similarly, the first episode in the "make money" story involved unsuccessful stock-market investments and the second episode involved successful real-estate investments. Each episode comprised four basic actions followed by an outcome; these statements alone made up the "Short" episode. By interspersing five more filler actions (statements) into a short episode, it was converted into a "Long" episode. Thus, our "Short" episode refers to four actions plus an outcome, whereas our "Long" episode refers to the same propositions but with five more filler actions put into the narrative (nine actions in total). The example below is the first episode for the "Job Seeking" story; sentences in parentheses were the fillers added to make the longer first episode.

> "...First, David went to a local employment agency. (At the agency he had to pay a fee), (then fill out a veritable mountain of forms.) (Finally, though, he was able to look through their job file.) He made a list of the job openings he was interested in. (The next day he began to visit the prospective employers.) The first place he visited said he was not qualified for the job. The second place said that they had no job, but were expecting to have one in the future. (The third place said the job was already filled). [OUTCOME] Finally, David decided that this was no way to get a job. [BEGIN EPISODE—2] Therefore, he decided to ask his acquaintances about jobs..."

Four versions of each story were constructed, consisting of a long or short first episode followed by a long or short second episode. A fourth of the subjects (eight) read each version of a story. The Short–Short stories were 150–200 words in length and the Long–Long stories were 250–300 words in length.

For each story, 16 recognition memory test items were composed. There were eight test items per episode. Four were verbatim "True" sentences: in fact, they were the four action statements in the Short version of that episode. They were printed in random order mixed in with four False lures, which were sentences expressing plausible events that could have happened to the actors in the episode but did not. The 16 test items for each

story were blocked together with their story title (e.g., JOBS).

Procedure. The subjects participated in this experiment in parallel with another experiment. They first read for 2 minutes some brief descriptions of stereotyped activities (eg., making coffee) that were unrelated to the stories of this experiment. They then received a four-page booklet containing the four stories described above. They were given 6 minutes to read the four stories, and warned after 5 minutes that they had 1 minute remaining. They were told to "read carefully in order to answer questions about the stories later." Across different subjects each story was presented in all four versions (Long or Short episode-1 followed by Long or Short episode-2). Over the four stories, each subject read one exemplar of a Long–Long story, a Long–Short one, a Short–Long one, and a Short–Short one. The order of story versions followed a Latin-Square with four sequences each sequence being read by a different group of eight subjects.

After the subjects had read the four stories. they had a free recall test and then a recognition memory test over the stereotyped activity scripts (eg., making coffee) they had read at the beginning of the experimental session. These tests took about 20 minutes.

Then the subjects had a recall test followed by a recognition test over the four stories of this experiment. Recall was cued by the title of each story and the subjects wrote their recalls. They were asked to recall as close to verbatim as possible, but to write down anything they could remember. They had as long as they wished to complete their recalls of each story; they typically took 2 or 3 minutes. After completing the four recalls the subjects were given the recognition memory test. This test contained 64 single sentences, four pages of 16 per story. The subjects were to rate each sentence on a 7-point scale their confidence that exactly that sentence had appeared in the text they had read; a rating of 1 denoted that they were certain they had not read that item, and 7 denoted they were certain they had read it. They wrote their recognition rating beside each test sentence. After the recognition test, they were debriefed and dismissed.

Results and Discussion

Recall. We are interested primarily in the recall of the basic actions and outcomes common to the Long and Short episodes. Each protocol was scored on a gist or substance basis for presence or absence of the basic actions and outcome from each episode of each story. Recall of these actions and outcomes tended to be "all-or-none" so we had no problems scoring marginal protocols.

The data of primary interest are summarized in Table 1 which reports recall percentages for the four basic actions in each episode for the four story versions. The chunk-independence hypothesis predicts that recall would not vary with the size of the other episode. The column means in Table 1 support this: recall of the basic actions in one episode was nearly the same whether the other episode was Long (83%) or Short (85%). Our primary prediction has therefore been confirmed by the recall results.

Table 1 reveals a significant *positive* effect of episode length upon recall of the basic actions

TABLE 1

PERCENTAGES OF RECALL OF BASIC ACTIONS RELATED TO THE LENGTH OF THE SAME OR THE OTHER EPISODE IN EXPERIMENT 1

	Other episode		
Same episode	Short	Long	Mean
Short	81	77	79
Long	88	89	89
Mean	85	83	

in that episode. Recall percentages are 79% for actions within Short episodes versus 89% for the same actions within Long episodes. These differ significantly whether the analysis-of-variance model considers story materials (episodes) as a fixed effect, $F(1,28)=15.68$, $p<.001$) or as a random effect, $F'(1,9)=5.38$, $p<.05$. Thus, the Long versus Short effect on recall can be generalized over subjects and over materials. We will return to discussion of this effect after we have examined other aspects of the data.

The results in Table 1 are for the four basic actions common to the Short and Long episodes. We tabulated separately recall of the outcome statements which closed out the episodes in the stories. As Mandler and Johnson (1977) had found before us, outcomes proved to be highly recalled, averaging 96%. Mean recall percentages varied from 94 to 97% for the several conditions, with no significant differences among the four story conditions. Thus, as the subjective "main point" of the episode, the outcome is practically always recalled.

Recognition memory. The results from the recognition memory test can be briefly summarized: there were no significant differences among conditions in hit rate (correct recognition of True sentences) or in false positive rate (accepting False sentences as Old). On the 7-point rating scale for recognition, the stated basic actions received average ratings varying from 5.24 to 5.39 for the four story conditions, and these do not differ significantly. Similarly, average ratings of False test sentences varied from 1.17 to 1.30, with no significant differences among conditions.

These results were expected. They follow the principle that recognition memory tests are unlikely to show effects of material organization (see Kintsch, 1977a). Organization primarily affects memory tests that involve large "retrieval" components, in which the person must produce the material from his own resources. Recognition tests usually minimize this component of remembering.

Experiment 2

We wanted to determine what characteristic of our materials in Experiment 1 led to the superior recall of the target actions in the long versions of the episodes. Although we had attempted to divide the episode actions into matched target and filler groups, we needed to check whether some bias had operated when we made the division. Hence, we collected "importance" ratings for both the target and filler actions.

Method

Subjects. The subjects were 12 Stanford undergraduates either fulfilling a requirement for their Introductory Psychology class or receiving payment.

Materials. The four stories used were the Long–Long versions from Experiment 1. Thus each of the stories had two episodes; each episode had nine actions (the four target actions and five filler actions from before) and an outcome. Numbers were inserted in the story texts to mark these 20 critical statements.

Procedure. The subjects were instructed to rate each of the numbered statements according to how "important" they were to the point or message of the story. They rated the statements on a 1 to 7 scale, with 1 meaning "very unimportant" and 7 meaning "very important." Each subject rated all statements in all four stories.

Results and Discussion

The ratings revealed that the filler actions were indeed less important than the target actions. The fillers received an average rating of 3.26, whereas the target actions received an average rating of 4.31. These ratings differ reliably, both with episodes and actions considered as a fixed effect, $F(1,11)=26.16$, $p<.01$, and as a random effect, $F'(1,18)=14.72$, $p<.01$. The outcomes were the most important of all: they were given an average rating of 5.36. This rating

was significantly greater than the target actions rating (4.31) both with episodes considered as a fixed effect, $F(1, 11) = 11.68$, $p < .01$, and as a random effect, $F'(1, 13) = 9.61$, $p < .01$.

Since the target actions were rated as more important than the filler actions, we expected them to be more salient in the memory representation of the story. Therefore, we would expect the target actions to be remembered better than the filler actions. Hence we reanalyzed the recall results in Experiment 1 for the long episodes, those containing both target and filler actions. We found, as expected, that the target actions were recalled significantly better (89%) than the filler actions (61%). These differences were reliable, with stories as a fixed effect, $F(1, 28) = 257$, $p < .001$, and with stories as a random effect, $F'(1, 17) = 39$, $p < .001$.

Since the outcomes were rated as more important than both the target and filler actions, we expected them to be the most salient in the memory representation of the story. The recall results in Experiment 1 showed that the outcomes were recalled significantly better (96%) than the target actions (84%) both with episodes as a fixed effect, $F(1, 28) = 14.85$, $p < .01$, and as a random effect, $F'(1, 10) = 5.21$, $p < .05$.

Unfortunately, we could not compare the outcomes, target actions, and filler actions in recognition memory because the Experiment 1 recognition test only contained the target actions.

Experiment 3

Next, we inquired into the reasons for why adding unimportant constituent statements might increase the recall of important constituent statements. We thought the relationship between the target and filler actions in our stories might be relevant. Examining the importance ratings for the individual target actions, we discovered one type of statement that always received a high importance ratings (i.e., at least 5 on our 7 point scale). These highly important actions were not only one of the actions in their episode, but could also serve as a summary of their episode. For instance, in the example episode contained under Method 3 of Experiment 1, the statement "David went to a local employment agency" is the first action in the episode; but notice that it together with the outcome also serves to summarize the entire episode ("David went to a local employment agency. He decided this was no way to get a job.")

Summarizing story actions of this kind will be called "superordinate" actions, and actions that are further specification of them will be called "subordinate" actions. The subordinate actions are of less importance in the story than the superordinate actions. We hypothesized that the more supporting subordinate actions subsumed under a superordinate action, the better should be the person's recall of that superordinate action. To test this hypothesis, we conducted another experiment in which we varied the number of subordinate actions and examined the effect on the free recall of the superordinate actions.

Method

Subjects. The subjects were 32 Stanford undergraduates either fulfilling a service requirement for their Introductory Psychology class or receiving payment.

Materials. The stories were revised versions of the four stories used in Experiment 1. The new versions each had two episodes, each with a clearly superordinate action along with six subordinate actions. Three of the subordinate actions were targets and three were "fillers." Each episode had a "Short" version which contained the superordinate action together with the three target subordinate action, and a "Long" version which added three filler subordinate actions to the "Short" version. Each episode also had an outcome statement. Thus combining the Long and Short versions of the episodes, each two-episode story had four versions: a Short–Short, Short–Long, Long–Short, and Long–Long version.

Procedure. The subjects were given a booklet containing the four stories and then given 5 minutes to read them. They were warned after 4 minutes that they had 1 minute remaining. After reading the main stories, the subjects received an intervening task. They were given a set of stories unrelated to the experimental stories and they wrote summaries of these stories for 20 minutes. After this intervening task, they received a memory test. They were cued with the titles of the four stories and they wrote their recalls. They were asked to recall as close to verbatim as possible, but to write down anything they could remember.

Each story was presented in all four versions (i.e., Short–Short, Short–Long, Long–Short, and Long–Long) across different subjects. Over the four stories, each subject read one example of each episode-length combination.

Results and Discussion

We were primarily interested in the recall of the superordinate action and how that relates to the number of subordinate actions in the same episode. As the rightmost column of Table 2 shows, 80% of the superordinate actions were recalled when the episode contained three subordinates, but their recall accuracy jumped to an impressive 98% when the constituent contained six subordinate actions. This difference is significant considering episodes as a fixed effect, $F(1,28)=23$, $p<.001$, or as a random effect, $F'(1,36)=6.3$, $p<.05$. Hence as predicted, the more subordinate actions subsumed under an action, the greater was the probability that the superordinate action would be recalled.

This experiment also provided an opportunity to replicate the episode-recall independence found in Experiment 1. We examined the recall accuracy of three different kinds of episode statements the superordinate actions, the target subordinate actions, and the outcome statements and noted how these related to the number of subordinate actions in the same and the other episodes.

Table 2 gave the results for the superordinate actions. As noted before, an increase in the number of subordinate actions in the same episode produced a highly significant increase in the recall probability of the superordinate action. However, Table 2 also shows a slight increase in superordinate action recall as the number of subordinate actions in the other episode increases. With three subordinates in the other episode, 86% of the superordinate actions were recalled; but with six subordinates in the other episode, 92% of the superordinate actions were recalled. This 6% difference is marginally significant with episodes as a fixed effect, $F(1,28)=4.57$, $p<.05$; but was far from significant with episodes as a random effect, $F'(1,4)=1.6$, $p<.20$. Hence, the number of subordinate actions in another episode in a story may have a slight effect on recall of superordinate

TABLE 2

RECALL PERCENTAGE OF SUPERORDINATE ACTIONS AS A FUNCTION OF THE NUMBER OF SUBORDINATE ACTIONS IN THE SAME OR THE OTHER EPISODE (EXPERIMENT 3)

	Other episode		
Same episode	Three subordinates	Six subordinates	Mean
Three subordinates	75	84	80
Six subordinates	97	100	98
Mean	86	92	

TABLE 3

RECALL PERCENTAGE OF TARGET SUBORDINATE ACTIONS AS A FUNCTION OF THE NUMBER OF SUBORDINATE ACTIONS IN THE SAME OR THE OTHER EPISODE (EXPERIMENT 3)

	Other episode		
Same episode	Three subordinates	Six subordinates	Mean
Three subordinates	61	59	60
Six subordinates	70	67	69
Mean	66	63	

actions, but the magnitude of this effect (6%) is much smaller than the effect of the number of subordinates in the same episode (18%), $F(1,28)=15$, $p<.001$, $F'(1,20)=5.12$, $p<.05$. Since the two episodes were part of the same story involving the same character and the same overall goal, we should not be surprised to find at least a slight dependency between their recall. The important fact is the relative magnitudes of the "list-length" effect within versus between episodes.

Table 3 presents the recall results for the target subordinate actions. Here, the number of subordinates in the other episode had no significant effect on recall (66% for short versus 63% for long), but the number of subordinates in the same episode has a positive effect. Thus, when an episode contains the three target subordinate actions alone, 60% of these actions were recalled; but when the three filler actions were added to the target actions (making a total of six subordinate actions), 69% of these target actions were recalled. This difference is significant with episodes as a fixed effect, $F(1,28)=6.25$, $p<.05$, and marginally significant with episodes as a random effect, $F'(1,20)=4.16$, $.05<p<.10$. Hence, adding the filler actions in the same episode facilitated recall of the target subordinate actions, but adding the filler actions in the other episode had no effect.

As with Experiment 1, neither the length of the same episode nor the length of the other episode had an effect on the recall of the outcome statements, whose average recall accuracy was 91%.

To summarize, then, we found a large increase in the recall of superordinate actions as we increased the number of subordinate actions in the same episode. This was as predicted. It suggests that adding less important actions to an episode increases the recall of more important actions primarily because subordinate elaborations facilitate recall of superordinate actions. However, we also found that adding filler subordinate actions slightly increased recall of target subordinate actions, suggesting some factors beyond subordinate–superordinate facilitation. What these other factors might be is a question for future research. This experiment also replicated the results of Experiment 1 with both superordinate and target subordinate actions by finding that the number of subordinate actions in the other episode had either no effect or a very slight effect on recall.

GENERAL DISCUSSION

The experiments have yielded three basic results. First, we found that goal-oriented episodes in stories are stored as separate chunks in the memory representation of the story. We demonstrated this separation by showing that recall of the actions in an episode was affected largely by the length of that episode but hardly at all by the length of the other episode. Second, we found that adding

related though subordinate actions to a story episode increased the recall probability of the important statements in the episode. In particular, the recall probability of a general superordinate action in a story increased with the number of subordinate actions that further specify it.

In line with our conclusions, recent papers by Mandler (1978) and Glenn (1978) provide further evidence that goal-oriented episodes act like separate chunks in memory. Mandler's subjects read stories in which actions from two episodes were interleaved; in free recall, her subjects tended to cluster the actions into their appropriate episodes. Children reorganized the interleaved stories into episodes even more than adults. Unfortunately, Mandler's story episodes had not only separate goals, but also different characters. Hence, the critical clustering factor could be the character (agent of the actions) rather than the episode structures. In contrast, in our experimental stories, the only basis for clustering the actions is that those in an episode form a sustained attempt to attain the goal of the story.

Glenn found that when her subjects read short elliptical story episodes, they tended to elaborate them during recall. Thus the subjects added statements during recall to make the episodes more coherent. The likelihood of these additions depended only on the episode length, not on the total length of the story. Hence these results also indicate that the episode is an organizational unit in memory for stories.

The finding that episodes act as separate chunks in memory provides evidence for the general class of "story grammars," and for the "macrostructure" theories of van Dijk (1977), Kintsch (1977b), and Schank and Abelson (1977); but provides evidence against the theories of Kintsch (1974), Schank (1975), Frederiksen (1972, 1975a, 1975b), and Frederiksen (1975c, 1977).

In our two recall experiments, adding more story statements either led to better recall of the other statements or had no effect on the recall of the other statements. There must surely be limits on this generalization. That is, one could probably compose texts for which adding more story statements will interfere with the recall of the original statements. Exploring how memory varies when more statements of various kinds are added to a text seems like a fruitful area for future research.

REFERENCES

BOWER, G. H. Chunks as interference units in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 610–613.

FREDERIKSEN, C. H. Effects of task-induced cognitive operations on comprehension and memory processes. In R. O. Freedle and J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston, 1972.

FREDERIKSEN, C. H. Acquisition of semantic information from discourse: Effects of repeated exposures. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 158–169. (a)

FREDERIKSEN, C. H. Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 1975, **7**, 139–166. (b)

FREDERIKSEN, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, **7**, 371–458. (c)

FREDERIKSEN, C. H. Semantic processing units in understanding text. In R. O. Freedle (Ed.), *Discourse production and comprehension*. Norwood, N.J.: Ablex, 1977.

GLENN, C. G. The role of episodic structure and of story length in children's recall of simple stories. *Journal of Verbal Learning and Verbal Behavior*, 1978, **17**, 229–247.

JOHNSON, N. F. The role of chunking and organization in the process of recall. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 1970. Vol. 4.

KINTSCH, W. *The representation of meaning in memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.

KINTSCH, W. *Memory and cognition*. New York: Wiley, 1977. (a)

KINTSCH, W. On comprehending stories. In M. A. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977. (b)

KINTSCH, W., & VAN DIJK, T. A. Toward a model of text comprehension and production. *Psychological Review*, 1978, **85**, 363–394.

MANDLER, J. M. A code in the node: The use of a story schema in retrieval. *Discourse Processes*, 1978, **1**, 14–35.

MANDLER, J. M., & JOHNSON, N. S. Remembrances of things parsed: Story structure and recall. *Cognitive Psychology*, 1977, **9**, 111–151.

MANDLER, G. Organization and Memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 1967. Vol. 1.

MEYER, B. J. F. *The organization of prose and its effect on memory*. Amsterdam: North-Holland, 1975.

RUMELHART, D. E. Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.

RUMELHART, D. E. Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

SCHANK, R. C. The structure of episodes in memory. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.

SCHANK, R. C., & ABELSON, R. P. *Scripts, plans, goals, and understanding*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

STEIN, N. L., & GLENN, C. G. An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *Discourse Processing*. Norwood, N.J.: Ablex, in press.

THORNDYKE, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, **9**, 77–110.

TULVING, E., & PEARLSTONE, Z. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 381–391.

VAN DIJK, T. A. Semantic macro-structures and knowledge frames in discourse comprehension. In M. A. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

(Received June 12, 1978)