Younger and Older Adults’ Use of Mental Models as Representations for Text Materials

Gabriel A. Radvansky, Linda D. Gerard, and Rose T. Zacks
Michigan State University

Lynn Hasher
Duke University

Recent work suggests that formation and use of mental models (representations of situations described) is an integral part of discourse comprehension. In an experiment comparing younger and older adults on this aspect of text comprehension, subjects heard readings of a list of sentences and took a forced-choice recognition test. The test contained 2 types of distractors with an equal degree of verbatim similarity to the target sentences. One type described the same situation as the corresponding target sentence; the other did not. If mental models are an integral part of text representation formed at encoding, then a large number of confusions of the first, but not the second, type of distractor with the target sentence would be obtained. Younger and older adults showed this pattern to equal degrees. These data are consistent with those indicating that linguistic competence remains stable over the adult years (cf. Light, 1988).

A major focus of recent empirical and theoretical work on text processing has been the role of mental or situational models in text comprehension (e.g., Garnham, 1981; Johnson-Laird, 1983; van Dijk & Kintsch, 1983). Theories invoking this concept argue that in understanding text, people construct not only a representation of the text itself but also one of the situation described by the text. The situational model is a level of representation that transcends the verbal–linguistic surface and propositional levels described by earlier models (most particularly, those of Kintsch, 1974, and Kintsch & van Dijk, 1978; cf. Perrig & Kintsch, 1985). Its construction is the result of an integration of text information with general background knowledge about the referent situation, including the likely characteristics of relevant people, objects, events, and environments. The result is a representation that elaborates on the text information to frequently include unstated pragmatic, perceptual, and spatial–temporal details. The mental model is assumed to be an essential component of a text representation because “it guides interpretation of the text, controls inference making, and influences the extent to which the text is judged coherent” (Glenberg, Meyer, & Lindem, 1987, p. 69).

Although the concept of mental models admittedly lacks a precise definition (cf. Glenberg et al., 1987), the central feature of a mental model is that it is a representation of what the text is about and not of the text itself. In particular, it represents a single state of affairs—a single situation—rather than being a general knowledge structure, such as a script (Schank & Abelson, 1977) or a schema (Alba & Hasher, 1983), although these knowledge structures may be used in the creation of mental models. This aspect of a mental model is readily illustrated by the Bransford, Barclay, and Franks (1972) study on constructive representations of text. In that research, subjects given a sentence-recognition test had difficulty distinguishing between a presented sentence and one congruent with the same mental model, whereas they easily distinguished between a presented sentence and a foil inconsistent with the target sentence’s mental model. For example, subjects made frequent errors in recalling whether they had been presented with the sentence “Three turtles rested on a floating log, and a fish swam beneath them” or the sentence “Three turtles rested on a floating log, and a fish swam beneath it.” Although these two sentences are very similar in wording and propositional content, these factors are not the primary cause of the high level of confusion. Instead, the important factor is the similarity of the situations they describe: In particular, the two sentences imply the same relative locations of the turtles, the log, and the fish. The evidence for this conclusion comes from the relatively good performance on sentence pairs in which the two sentences implicated different mental models while being as similar to each other at the verbal–linguistic and propositional levels as the same-model pairs. For example, subjects made few errors in discriminating between the sentence “Three turtles rested beside a floating log, and the fish swam beneath it” and the sentence “Three turtles rested beside a floating log, and the fish swam beneath them.” Presumably, the improvement resulted because the two “beside” versions evoke different mental models of the spatial relations among the turtles, the log, and the fish.

To date, the new emphasis of text-processing theorists on mental models has not been reflected in the cognitive gerontology literature. This article represents an attempt to begin to correct this omission.

In the absence of any data on older adults’ formation and use of mental models, making predictions about whether there will be any age differences in this aspect of discourse processing is

---

This research was supported by National Institute on Aging Grant AG04306. We acknowledge the able assistance of Roderick McFarland.

Correspondence concerning this article should be addressed to Gabriel A. Radvansky or Rose T. Zacks, Department of Psychology, Michigan State University, East Lansing, Michigan 48824.
hard. However, the major conclusions from Light's (1988) recent comprehensive review of research on linguistic functioning in old age provides some basis for speculation. According to Light, the existing literature indicates little or no change in linguistic competence with increasing age. Older adults, for example, show undiminished sensitivity to such subtle semantic and syntactic distinctions as differences in the implicit causality of verbs: Like young people, older adults interpret the ambiguous pronoun he to refer to John in "John apologized to Bill because he . . ." and to Bill in "John blamed Bill because he . . ." (Light & Capps, 1983).

Of course, linguistic performance frequently is found to show age declines. Many have argued that these performance declines seem to be tied to high demands on limited cognitive resources, including working memory. For example, there are exaggerated age deficits in the production and comprehension of syntactic structures, such as left-branching sentences, which require the on-line storage of considerable information in working memory (cf. Kemper, 1988).

Generalizing from Light's (1988) analysis, we reasoned that a prediction about whether there will be age deficits in the construction and use of mental models would have to take account of the on-line processing demands of the experimental task. We thus began our investigations into the formation of mental models using texts that make relatively small demands on processing capacity. Specifically, the materials were relatively short sentences with fairly simple syntactic structures. Thus, the opportunity to show preserved competence was provided, and a finding of age invariance would not be surprising.

To investigate age trends in the use of mental models in sentence comprehension, we used a modification of a paradigm developed by Garnham (1981). (As we explain later, Garnham's procedure is conceptually analogous to that of Bransford et al., 1972.) Garnham compared pairs of sentences that were likely to evoke the same mental model with pairs that were likely to evoke different mental models. As an example of the first type of sentence pair, consider the following two sentences:

A. The girl was given a complete pedicure at the podiatrist's.
B. The girl was given a complete pedicure by the podiatrist.

Even though these two sentences differ in their exact wording and in their propositional content (Sentence A describes where the girl received the pedicure, and Sentence B describes who gave it to her), they are both plausible descriptions of the same situation and therefore are likely to evoke the same mental model—that is, of the girl receiving a pedicure from the podiatrist at the podiatry office. Garnham (1981) found that target-distractor pairs like Sentences A and B produced a high confusion rate in recognition.

In contrast, consider the next pair of sentences:

C. The girl had her handbag stolen at the podiatrist's.
D. The girl had her handbag stolen by the podiatrist.

As is the case for the first pair, Sentences C and D differ in their exact wording and in their propositional content (Sentence C describes where the girl had her handbag stolen, and Sentence D describes who stole it). Moreover, these differences exactly parallel those between Sentences A and B. However, unlike the A–B pair, the C–D pair is not normally consistent with a single state of affairs that would be represented by a single mental model. Because a professional is considered unlikely to steal the handbag of a client who comes to his or her place of business, the mental model of Sentence C differs from that of Sentence D, which directly states that the podiatrist was the robber. In confirmation of this analysis, Garnham (1981) found that subjects made few confusion errors with sentence pairs like C and D.

As can be seen in the previous examples, mental models of text are "organized around representations of events, rather than the linguistic expressions describing those events" (Garnham, 1981, p. 561). This allows sentences having parallel differences in wording and in propositional content to sometimes evoke a single mental model. Young people's recognition performance shows sensitivity to this difference. In this study, we examine whether the same is true for older adults.

Method

Sixteen older and 16 younger adults participated in this study. Two additional older adults had their data replaced because of an exceedingly large number of errors during the recognition test. The older adults ranged in age from 62 to 82 years (M = 71.5, SD = 6.4) and had 11–18 years (M = 13.3, SD = 2.3) of formal education. They were recruited from various area retirement residential complexes and senior citizen organizations. They were paid $4 for their participation. All were self-dependent and all reported themselves to be in relatively good health for their age. The younger subjects ranged in age from 18 to 26 years (M = 19.6, SD = 2.1) and had 12–19 years (M = 13.1, SD = 1.8) of formal education. These individuals were members of the Michigan State University subject pool participating to fulfill a course requirement.

All subjects were administered two individual-differences measures, the vocabulary subtest of the Wechsler Adult Intelligence Scale–Revised (WAIS-R) as a measure of verbal ability and Daneman and Carpenter's (1980) sentence-span test as a measure of working memory capacity. On the vocabulary test, the older group, with a mean raw score of 53.2 (out of 70; M = 8.3, SD = 2.6) performed better than the younger group, which had a mean raw score of 45.2 (SD = 6.3), t(30) = 3.08, p < .01. On the sentence-span test, by contrast, the average performance of the older subjects (M = 2.6, SD = 0.6) was slightly, but not significantly, lower than that of the younger subjects (M = 3.0, SD = 0.7), t(30) = 1.81, p > .05. Our samples are not unusual on either measure. Unless explicit attempts are made to equate verbal ability, the older subjects participating in cognitive gerontology research frequently have higher verbal ability scores than do the younger adults (e.g., Gick, Craik, & Morris, 1988; Tun, 1989). In addition, older adults typically achieve lower scores than do younger adults on the Daneman and Carpenter measure, although

---

1 The Daneman and Carpenter (1980) sentence-span test of working memory capacity requires subjects to read aloud a list of sentences. Sentences range from 13 to 16 words in length, and each is printed on a separate index card. When reaching a designated point (a blank card), the subject is to retrieve the last word from the most recent set of sentences. The set sizes range from two to five sentences. Individual subjects' memory-span scores reflect the largest set size in which they are able to recall all of the final words.
frequently, as in our samples, the age difference is not significant (e.g., Hartley, 1988; Light & Anderson, 1985).

Materials and Design

The 24 sets of test sentences were adapted from Garnham (1981). Each set consisted of four related sentences arranged in two pairs. The two sentences in one pair were designed to elicit the same mental model (confusable pair); those in the other were designed to elicit different mental models (nonconfusable pair). The specific features of these materials were as follows: All four sentences in a set shared the same subject (e.g., the girl). The two sentences in the confusable pair differed only in the final prepositional phrase, as did the two sentences in the nonconfusable pair. For half the sets, the two versions of this phrase were at versus by phrases (e.g., “at the podiatrist’s” vs. “by the podiatrist”); for the other half, the two versions were in versus from phrases (e.g., “in the optician’s” vs. “from the optician”). Sentences A-D, presented earlier, formed 1 of the sets used in the study. Another set is the following:

Confusable pair
E. The judge got his contact lenses in the optician’s.
F. The judge got his contact lenses from the optician.

Nonconfusable pair
G. The judge answered a telephone call in the optician’s.
H. The judge answered a telephone call from the optician.

A separate group of 14 young adults validated these materials. These individuals were given rating forms containing pairs of sentences. For each pair, they were asked to consider the situations described by the two sentences and to rate the similarity of those two situations. A 7-point scale was used, ranging from very similar situations (1) to very dissimilar situations (7). Subjects were instructed to circle the appropriate number. Each pair of sentences was either the confusable or the nonconfusable pair from a pool of four-sentence sets. There were two versions of the rating form so that each subject saw only one pair from each set, and each rating form contained half confusable and half nonconfusable pairs. For the 24 sentence sets selected for the experiment, the mean situational similarity ratings for the confusable and nonconfusable pairs were 2.51 and 5.35, respectively.

In addition to the experimental sentences, there were 12 filler sentences that had a structure similar to the experimental sentences but with prepositional phrases containing the word under or over.

Four presentation lists were generated, each containing the 12 filler sentences plus 1 sentence (a different one for each list) from each of the 24 experimental sets. Within each list, half of the experimental sentences were from the confusable pair, and half were from the nonconfusable pair. Also, half were the at-by type and half were the in-from type. The presentation order of each list was random except that the first 2 and last 2 sentences were always fillers. Four subjects in each age group received each of the lists.

The lists were recorded on cassette tape by an individual who was unfamiliar with the experimental hypotheses. He used a normal speaking voice. The intersentence interval was approximately 3 s.

The same four-alternative, forced-choice recognition test was used for all subjects. It contained the 24 sets of experimental sentences. The subjects had to decide, guessing if unsure, which one of the four sentences in each set had been heard on the tape.

As a final task, the participants were asked to perform the same rating task as the subjects who had validated the experimental materials. These ratings provided further validation of the materials, especially for older adults, who had not been represented in the validation group. Two rating forms were generated, each containing 24 pairs of sentences representing all the experimental sets. As a result of the way the forms were constructed, half of the pairs contained sentences heard earlier, and half did not. The 24 pairs on a form included 12 each of the confusable and nonconfusable types, and within each of these subsets, half were of the at-by and half of the in-from type. The pairs and the sentences within pairs were listed in a random order on the rating form. The rating scale was the same as that used for the validation group.

Procedure

Each subject was tested individually. Subjects were first administered the WAIS-R vocabulary subtest and the Daneman and Carpenter (1980) sentence-span test. Next, the tape of sentences was played on a portable tape recorder with the volume set at a moderate level. No subjects reported difficulty in hearing the tape. As a cover task, subjects rated the ease of imagining the situation described in each sentence on a 7-point scale that ranged from very easy to imagine (1) to very hard to imagine (7). A rating sheet was provided, and subjects circled the appropriate number for each sentence. This task of rating the ease of imagining the situations described by the sentences may have biased the subjects toward constructing mental models in a less-than-spontaneous manner compared with some other task, such as rating pleasantness. However, considering that comprehension means that people are trying to figure out what the sentence is about (in this case, to imagine what is being described in the sentence), this task does not appear inappropriate. The subjects were not informed of the upcoming tests because we wanted to discourage attempts at verbatim memorization of the sentences.

After the presentation of the sentences, subjects performed a distractor task consisting of a paper-and-pencil test of 48 three-digit addition problems (e.g., 385 + 298). Subjects were asked to work at this task (a) until they had completed all 48 problems or (b) for a maximum of 10 min. The surprise recognition test followed immediately after the completion of the distractor task. The recognition test items were presented in a three-page booklet. The four sentences in a set were presented together, with the sets clearly separated from each other. Both the order of the sentence sets and of the sentences within each set were random. A space was provided for each set for subjects to indicate which of the four sentences had been presented earlier. After the recognition test, subjects were presented with the situation-similarity rating form. Both the recognition test and the similarity ratings were self-paced with no time limit. Each session lasted approximately 30 min.

Results

Cover-Task Ratings

Subject ratings for the ease of imagining the situations described by the sentences averaged 2.02 (SD = 0.66). Confusable sentences were rated as easier to imagine than nonconfusable sentences, with mean ratings of 1.59 (SD = 0.44) and 2.45 (SD = 0.57), respectively, t(46) = 5.77, p < .001. The lower mean rating for the confusable sentences can be attributed to their greater congruency with general knowledge structures, such as scripts (see the Discussion section), aiding in sentence-representation formation.

Rated Similarity of Sentence Pairs

Both age groups rated the two sentences in confusable pairs as having greater situational similarity than those in nonconfusable pairs. The ratings for the younger adults were 2.92 and 5.14

\(^2\)A complete list of all of the materials used in the study is available from the authors on request.
Recall Test Performance

For the confusable and nonconfusable pairs, respectively. The comparable ratings for the older adults were 2.24 and 4.94. The difference between confusable and nonconfusable pairs was significant, \( F(1, 30) = 127.06, p < .001, MS_e = 0.76 \). Neither the age effect nor the Age X Pair Type interaction was significant. Therefore, the distinction between confusable and nonconfusable pairs appears to be equally valid for the two age groups.

**Table 1**

**Percentages of Correct Responses and Errors on the Recognition Test**

<table>
<thead>
<tr>
<th>Age group</th>
<th>Confusable</th>
<th>Nonconfusable</th>
<th>Overall mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Younger adults</td>
<td>Correct responses</td>
<td>59.4</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td>Confusion errors</td>
<td>38.0</td>
<td>12.2</td>
</tr>
<tr>
<td></td>
<td>Nonconfusion errors</td>
<td>2.6</td>
<td>5.0</td>
</tr>
<tr>
<td>Older adults</td>
<td>Correct responses</td>
<td>56.2</td>
<td>12.0</td>
</tr>
<tr>
<td></td>
<td>Confusion errors</td>
<td>39.6</td>
<td>11.6</td>
</tr>
<tr>
<td></td>
<td>Nonconfusion errors</td>
<td>4.2</td>
<td>6.1</td>
</tr>
</tbody>
</table>

**Verbal Ability and Sentence-Span Differences**

Various analyses were performed to explore the relationship between the two individual-differences measures obtained and performance on the experimental task. Because of the small number of subjects in each age group and because of the restricted range in one of the individual-differences measures (the sentence-span test), the results of these analyses are only suggestive and are not discussed in detail. Correlational analyses were performed to compare WAIS-R vocabulary score, sentence span, and overall percentage correct on the recognition test. With age partialed out, the correlation between the vocabulary and sentence-span scores was .45; between vocabulary score and total correct recognition, it was .22; and between sentence span and total correct recognition, it was .10 (with 30 dfs, a correlation of .35 is required for significance at \( p = .05 \)). The low level of the last of these correlations is perhaps surprising, given Daneman and Carpenter’s (1980) data showing that their measure is a strong predictor of language comprehension. However, other studies (e.g., Hartley, 1988; Light & Anderson, 1985) have shown weaker relationships between sentence span and performance on discourse-processing tasks, and as already noted, there was a restricted range of sentence-span scores among our subjects.

Next, the correlations for each age group were considered separately. For the younger and older subjects, respectively, the correlations between the vocabulary and sentence-span scores were .33 and .53; between vocabulary score and total correct, they were -.24 and .49; and between sentence span and total correct, they were -.17 and .41 (with 14 dfs, a correlation of .50 is needed for significance at \( p = .05 \)). In contrast to those of the younger people, the correlations between these individual-
differences measures and the performance measure were positive and marginally significant for the older adults, suggesting that the older adults may be more influenced by these abilities.

Subjects were split into high- and low-score subgroups on the basis of a median split within age groups for the vocabulary measure. For the memory-span measure, they were divided into high and low groups by considering scores below 3 as low. The lower ability group of older adults tended to make more total errors than did the higher ability group of older adults, whereas the younger subjects did not show such a difference. For sentence span, high-scoring and low-scoring older adults made 31.7% and 42.8% errors, respectively, on the sentence-recognition test compared with 33.0% and 32.5%, respectively, for the younger adults. A similar pattern was obtained with the vocabulary scores, with 34.9% and 43.7% errors for the high-scoring and low-scoring older adults and 33.9% and 31.8% errors, respectively, for the young. However, two ANOVAs using these variables as between-subjects factors were not significant. (The small sample size made a combined analysis highly questionable.) Although far from conclusive by themselves, these data join other hints in the literature (e.g., Zacks, Hasher, Doren, Hamm, & Atting, 1987) that aging deficits might be reduced for high-ability older adults.

Discussion

The results suggest that older adults generate and use mental models much as young adults do when they process text materials. For both age groups, the findings replicate those of Garnham (1981). In particular, errors were more frequent when the presented sentence was consistent with the same mental model as one of distractors than when it was not, and in such cases, the errors were mainly choices of the confusion distractor. Beyond the slightly higher overall error rate of the older adults, which was magnified by decreased verbal ability or working memory span, it is striking that there was no indication of age differences in performance on our task. Therefore, the following description seems to apply to both age groups: When hearing a sentence, the subjects constructed a mental model of the situation described. During the recognition test, if there was more than one option consistent with the situation represented by the mental model, the subjects had greater difficulty selecting the correct sentence than if this was not true. This description implies that it is the mental model that is used to guide selection during the recognition task.

This result parallels the finding by Light and Anderson (1985) that both younger and older subjects showed similar patterns in the use of scripts. The important distinction here is that a script is a general knowledge structure of a stereotypical sequence of events (Schank & Abelson, 1977), whereas a mental model represents a single state of affairs (Johnson-Laird, 1983), possibly never encountered before. Although we prefer to interpret our results in terms of mental models, we note that the results can be explained in terms of scripts. The confusable sentences tend to be script conforming (e.g., it is more likely that there is a script for buying a mink coat from a furrier in the furrier's) but not the nonconfusable sentences (e.g., it is less likely that there is a script for receiving a telegram from a furrier). Both sentences from the confusable pairs conform to the instantiated script, thus leading to a higher rate of confusion errors on the recognition test. However, both the script and mental model explanations seem to be getting at the same point; both older and younger adults comprehend sentences such that they are likely to make recognition errors when presented with sentences that describe the same situation. Whether this situation is represented by an instantiated script or a mental model is beyond the scope of this article. However, it is clear that not only are older adults as able as younger adults to use highly regularized knowledge structures during comprehension, but the older adults are also just as able to generate and use representations of novel situations.

Of course, our data are only suggestive. It cannot be concluded from them that the use of mental models in text processing will be age invariant under all circumstances. In particular, for texts or presentation modes or both that place high demands on working memory during encoding, the mental models generated by older adults may be less fully articulated than those generated by young adults (cf. Light, 1988). Also, older adults may have more difficulty when an initial mental model must be changed because an incorrect one was generated. (For a theoretical view that makes the latter prediction, see Hasher & Zacks, 1988.) Still, this study indicates that there are situations (e.g., the use of simple sentences in a task placing low demands on working memory) in which older and younger adults are similar in their formation of mental models of the presented texts.

References


Received February 22, 1989
Revision received August 14, 1989
Accepted August 17, 1989

Beutler, Levin, Tesser, and Miller Appointed

The Publications and Communications Board of the American Psychological Association announces the appointments of Larry E. Beutler, University of Arizona; Joel R. Levin, University of Wisconsin; Abraham Tesser, University of Georgia; and Norman Miller, University of Southern California, as editors of the *Journal of Consulting and Clinical Psychology*, the *Journal of Educational Psychology*, the Attitudes and Social Cognition section and the Interpersonal Relations and Group Processes section of the *Journal of Personality and Social Psychology*, respectively. As of January 1, 1990, manuscripts should be directed as follows:

- For *Educational* send manuscripts to Joel R. Levin, Department of Educational Psychology, University of Wisconsin, 1025 West Johnson Street, Madison, Wisconsin 53706.
- For *JPSP: Attitudes* send manuscripts to Abraham Tesser, Institute for Behavioral Research, University of Georgia, 548 Boyd Graduate Studies, D. W. Brooks Drive, Athens, Georgia 30602.
- For *JPSP: Interpersonal* send manuscripts to Norman Miller, Department of Psychology, Seeley G. Mudd Building, University of Southern California, University Park, Los Angeles, California 90089.