Situation Models and Aging

Gabriel A. Radvansky
University of Notre Dame

Rolf A. Zwaan
Florida State University

Jacqueline M. Curiel
University of Texas at San Antonio

David E. Copeland
University of Notre Dame

Younger and older adults were tested for their ability to process and retrieve information from texts. The authors focused on the construction and retrieval of situation models relative to other types of text representations. The results showed that during memory retrieval, younger adults showed superior memory for surface form and textbase knowledge (what the text was), whereas older adults had equivalent or superior memory for situation model information (what the text was about). The results also showed that during reading, older and younger adults were similar in their sensitivity to various aspects of the texts. Overall, these findings suggest that although there are age-related declines in the processing and memory for text-based information, for higher level representations, these abilities appear to be preserved. Several possibilities for why this is the case are discussed, including an in-depth consideration of one possibility that involves W. Kintsch's (1988) construction-integration model.

The aim of this article is to explore changes in comprehension and memory that occur as a result of the natural aging process. There is some agreement in theories of comprehension and memory that there are three basic levels of representation: the surface form, the propositional textbase, and the situation model (Johnson-Laird, 1983; Kintsch, 1998; Schmalhofer & Glavanov, 1986; van Dijk & Kintsch, 1983; Zwaan, 1994). The surface form corresponds to the actual words and syntax that were used. The textbase is an abstract representation of the information explicitly conveyed in the text but that is not tied to a specific external form. Finally, the situation model represents the events described by a text. It represents what the text is about, not the text itself. The creation of an accurate situation model is the main goal of comprehension and is what needs to be retained in memory if the knowledge is to be useful later. As such, it is important to understand how situation model use is influenced by the aging process.

Situation Models

A situation model represents a described situation (Johnson-Laird, 1983; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998) and serves as a mental simulation. Situation model use has been studied in a number of domains, including memory retrieval (Radvansky & Zacks, 1991), language comprehension (Zwaan, 1996), and logical reasoning (Johnson-Laird, Byrne, & Schaeeken, 1992). The focus here is on older adults' situation model use in memory and language.

There has been relatively little work on situation model use by older adults. The research that has been done has shown that aging has little effect. One of the first studies by Radvansky, Gerard, Zacks, and Hasher (1990) found that people made recognition decisions using situation models rather than a propositional representation. In that study, people overwhelmingly made errors of selecting a recognition choice that described the same situation as the originally heard sentence. Significantly, Radvansky et al. did not observe any differences in the pattern of errors of younger and older adults, suggesting that they use similar types of situation models.

Furthermore, Radvansky, Zacks, and Hasher (1996) reported that younger and older adults were similarly able to integrate information into situation models to reduce or eliminate interference during memory retrieval. Research by Morrow and colleagues (e.g., Morrow, Leirer, & Altieri, 1992; Morrow, Stine-Morrow, Leirer, Andrassy, & Kahn, 1997) has shown that the ability to update situation models during comprehension is largely unaffected by aging. Finally, Gilinsky and Judd (1994) reported that older and younger adults use situation models similarly when solving categorical reasoning problems. Although older adults performed more poorly overall, they did not differ in the ability to create and use situation models.

The lack of evidence that aging affects situation model use is surprising given the extensive evidence showing that older adults are compromised at lower levels of processing, such as syntactic processing (e.g., Kemper, 1987; Kemtes & Kemper, 1997; Light & Capps, 1986) and propositional textbase memory (e.g., Cohen, 1979; Dixon, Simon, Nowak, & Hultsch, 1982; Frieske & Park, 1979; Light & Capps, 1986).
Although older adults have difficulty with lower levels of processing, they are apparently able to compensate for this and gain an accurate understanding of the message being conveyed.

Before this issue can be pursued in detail, we first need to establish that relative to younger adults, older adults’ use of textbase representations is less effective and that their processing is dominated more by situation models relative to other levels of representation. It is possible that the studies that have failed to show an age effect in the ability to create and use situation models have been unable to reveal age-related deficits due to methodological problems. A major aim of the current study was to provide some resolution to this issue.

Aging and Comprehension

A study by Spilich (1983), although not presented from a situation model view, may provide some insight as to whether older adults’ processing is dominated more by situation models. In this study, people read narratives with the goal of later recalling them. A recognition test was then given with various memory probes: (a) verbatim, (b) syntactic change–semantic same, (c) syntactic same–semantic change, (d) situationally consistent but incorrect, and (e) semantically true but incorrect. The results showed that younger adults identified verbatim probes better than did older adults, suggesting that aging worsens surface level representations. Also, older adults were more likely to accept syntactic change–semantically same probes than were younger adults, but not syntactic same–semantic change probes. This is consistent with the idea that older adults focus more on what the text was about, rather than the text itself, consistent with an emphasis on situation models. Finally, older adults performed worse than the younger adults on situationally consistent but incorrect but not semantically true but incorrect probes. This is also consistent with the idea that situation models play a more dominant role in older adults’ processing. Although suggestive, Spilich’s study was not designed to tease apart the representations at all three levels.

More recently, Stine-Morrow, Loveless, and Soederberg (1996) studied aging effects at all three representational levels concurrently. People read texts, and reading times were collected. Significantly, aging effects were found at the surface and textbase levels but not at the situation model level (see also Miller & Stine-Morrow, 1998; Morrow et al., 1997). However, situation model use was based on a single measure: serial position. The idea was that when creating a model, the initial parts of the text are needed to establish its structure, making subsequent parts easier to integrate (Gernsbacher, 1990). Although we agree with this, serial position is a relatively shallow measure of situation model use because it is tied directly to the structure of the text itself rather than to the situation being described. In fact, in our analyses, serial position is an auxiliary variable rooted in the structure of the physical text itself rather than a situation index. In this light, the Stine-Morrow et al. study is limited in what it reveals about situation model use.

Finally, Adams (1991; Adams, Smith, Nyquist, & Perlmutter, 1997) has suggested that there are qualitative differences in younger and older adults’ text memory. Younger adults’ memories are rooted more in the text itself, whereas older adults’ are rooted more in interpretation. When people read texts, the recalls of younger adults were dominated more by information that was actually in the text, whereas those of older adults were dominated more by summary statements. Although this is consistent with our hypothesis, the method for distinguishing situation-oriented from textbase-oriented information was relatively weak. It does not provide a stable index of the degree to which people use textbase or situation model information in memory.

Methodological Overview

The current study used two experimental paradigms. One was developed by Schmalhofer and Glavanov (1986; see also Fletcher & Chrysler, 1990; Kintsch, Welsch, Schmalhofer & Zimmer, 1990; Zwaan, 1994) to separate out the influence of the surface form, textbase, and the situation model levels in memory performance. The other was a reading-time analysis developed by Zwaan (Zwaan, Magliano, & Graesser, 1995; Zwaan, Radvansky, Hiltiard, & Curiel, 1998) to assess the influence of various situation dimensions during comprehension.

The Schmalhofer and Glavanov (1986) paradigm, after reading a text, people are given a recognition test. There are four types of probes: (a) verbatim probes of the sentences that had appeared in the text, (b) paraphrase probes that contain the same propositions but that are expressed differently, (c) inference probes that were not mentioned but that are consistent with the described situation, and (d) incorrect probes. The ability to discriminate verbatim from paraphrase probes is an index of the surface representation. The difference between these is that one matches the original surface characteristics and the other does not. They are equivalent in how they map onto the textbase and situation model. The ability to discriminate between verbatim and paraphrase probes is based on the surface representation. Similarly, the ability to discriminate between paraphrase and inference probes is an index of the textbase. The difference between these is that one matches the original in propositional content, and the other does not. Also, both are inconsistent with the surface structure and consistent with the described situation. Finally, the ability to discriminate between inference and incorrect probes provides an index of the use of situation models. The difference between these is that one corresponds to the described situation, and the other does not. Both are inconsistent with the surface structure and textbase.

A' scores (after Donaldson, 1992), a nonparametric signal detection measure, were calculated as a discrimination measure for each of the above comparisons. For the surface form measure, verbatim was considered a hit, and paraphrases were considered a false alarm. For the textbase measure, paraphrases were considered hits, and inferences were considered false alarms. Finally, for the situation model measure, inferences were considered hits, and incorrects were considered false alarms. In addition, B' bias scores were also calculated as a measure of response bias using the same approach as the A' measure. However, it should be noted that the bias scores were not a measure of primary interest here.

In addition to the recognition measures, Zwaan and colleagues’ reading-time analysis derived from the event indexing model (Zwaan, 1999; Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995) was used. In this paradigm, reading times were submitted to a multiple-regression analysis that included a number of auxiliary and theoretical (situation model)
The third was
19) and absolute time (e.g., dates, in-
.10). However, it was significantly correlated
rs
<p>
(20, 1986,

<

P

adul
t
= 19.3) and had 12 to 20 years of formal education
M
= 3.99,
P
< .05. Two additional people, 1 from
each age group, were replaced for having exceptionally low vocabulary scores (less than 21 out of 40).

Materials

Four texts were used. These texts were 58 to 85 sentences long and described different historical events. These were the development of the marine chronometer in 18th-century England, the British Gunpowder Plot of 1605, the Australian Rum Rebellion of 1807, and the Dutch tulip craze of the 17th century. These were chosen because people were unlikely to have knowledge about them, yet they were readily understandable. The texts were derived from encyclopedias and similar sources. One text is presented in Appendix A. In addition, there was a practice text on the history of vegetarianism that was 34 sentences long.

For the recognition test, eight sentences were selected from each experimental text. Four memory probes were created based on each of the selected sentences. The first was the verbatim probe. This was an actual sentence from the text (e.g., “The plot bitterly intensified Protestant suspicions of Catholics”). The second was the paraphrase probe. This was a rewording of the sentence that retained the propositional content of the original (e.g., “The plot greatly heightened Protestant distrust of Catholics”). This was done using synonyms or altering word order. The third was the inference probe. This was information that was important to the description but was not explicitly mentioned (e.g., “The plot led to increased acts of persecution of Catholics”). Finally, the fourth was the incorrect probe. This was information that was not mentioned and was unlikely to be inferred. However, the information was globally consistent with the passage’s theme (e.g., “After the plot, donations to Protestant churches rose dramatically”).

Method

Participants

Forty-eight people were recruited for each of the two age groups. The older adults ranged from 61 to 96 years in age (M = 73.5) and had 10 to 20 years of formal education (M = 12.9). These people were recruited from a local senior citizens’ center, provided their own transportation to the University of Notre Dame, where the study was conducted, and were paid for their participation. The younger adults ranged from 18 to 26 years in age (M = 19.3) and had 12 to 20 years of formal education (M = 13.5). These people were either recruited from the participant pool at Notre Dame and were given partial class credit for their participation or were recruited through ads placed around campus and paid for their participation. There was no age difference in the number of years of education, r(94) = 1.56, p > .10. Older adults performed better on the Shipley vocabulary test (Zachary, 1986, M = 33.4) than younger adults, M = 30.6, r(94) = 3.37, p < .05. In contrast, using Daneman and Carpenter’s (1980) working-memory span test, younger adults performed better, M = 3.4, than older adults, M = 2.7, r(94) = 3.99, p < .05.1 Two additional people, 1 from

Experiment 1

The aim of Experiment 1 was to assess the influence of aging on multiple levels of representation concurrently. This was done using the same text and the same testing methodology to test each representational level. If the use of situation models is preserved with aging while there are declines at lower levels of processing, then one would expect no decline at the situation model level but expect declines at the lower levels of representation.

Consistent with prior studies using Schmalhofer and Glavanov’s (1986) paradigm, there were two instruction groups. Half of the people were told that they would be asked to summarize the text (summary group). This was designed to encourage people to create a more elaborate textbase representation. The others were told to acquire knowledge from the text and relate it to knowledge about other historical periods (knowledge acquisition group). This was designed to encourage the creation of more elaborate situation models. However, no effects of instruction type were found. In addition, it should be noted that we used history texts to make the use of Zwaan and colleagues’ (Zwaan, Migliano, & Graser, 1995; Zwaan, Radinsky, Hilliard, & Curiel, 1998) reading-time analysis more straightforward. Historical events contain many of the elements for which the event indexing model was developed—people, places, times, causes, and goals. Also, because history texts are more expository, the summary and knowledge acquisition instructions could be interpreted in a straightforward manner.

For the recognition test, eight sentences were selected from each experimental text. Four memory probes were created based on each of the selected sentences. The first was the verbatim probe. This was an actual sentence from the text (e.g., “The plot bitterly intensified Protestant suspicions of Catholics”). The second was the paraphrase probe. This was a rewording of the sentence that retained the propositional content of the original (e.g., “The plot greatly heightened Protestant distrust of Catholics”). This was done using synonyms or altering word order. The third was the inference probe. This was information that was important to the description but was not explicitly mentioned (e.g., “The plot led to increased acts of persecution of Catholics”). Finally, the fourth was the incorrect probe. This was information that was not mentioned and was unlikely to be inferred. However, the information was globally consistent with the passage’s theme (e.g., “After the plot, donations to Protestant churches rose dramatically”).

Procedure

At the beginning of the session, participants filled out a general-information form and then completed the vocabulary and working-memory span tests. After this, they read the practice and experimental texts. These were presented one clause at a time on a computer. The words appeared halfway down and left justified on the screen, in white on a black background, with the screen in 40-column mode. At the beginning of each text was a title (e.g., “The Rum Rebellion”), in yellow, to identify the topic. Participants advanced through the text by pressing the left button on a computer mouse. Reading times were recorded. The practice text was always presented first, but the order of the four experimental texts was randomized for each participant.

After reading all of the texts, the recognition test was given. The task was to indicate whether a sentence had been read earlier. Participants were warned that the sentences might be similar to the ones they read but contained some wording changes. The probes were blocked by story, with the title appearing before the probes. This was done so that participants knew to which passage the items referred. The practice story probes were always presented first, followed by the experimental story probes. The story order was the same as during reading. The order of the probes within

1 For the recognition test in both Experiments 1 and 2, working-memory capacity was not significantly correlated with surface form or model measures (all r ≤ .06, n = 10). However, it was significantly correlated with the textbase measure (r = .29 and r = .22, respectively). For the reading-time data, in Experiment 1, the only significant correlation was between span and number of new arguments (r = .24), whereas in Experiment 2, working-memory span was significantly correlated with mean frequency (r = -.19) and absolute time (e.g., dates, r = .20). These findings are generally consistent with the idea that Daneman and Carpenter’s (1980) reading-span task taps into the same processes involved in the construction of a propositional textbase.
a story block was randomized. Participants responded by pressing the left button on the mouse, marked with a Y, to indicate, “Yes, I did read this sentence.” or by pressing the right button marked with an N, to indicate, “No, I did not read this sentence.”

At the end of the recognition test, participants were given a questionnaire concerning their prior knowledge of the events described in the texts. This questionnaire listed the four topics in the experimental passages, along with a 7-point Likert-type scale that ranged from 0 to 6. Zero indicated no knowledge, and 6 indicated extensive knowledge. Participants responded by circling the appropriate number for each item.

**Analyses**

The recognition data were analyzed using A’ discrimination and B’ bias scores as described in the introduction. The reading-time means for each participant were submitted to a multiple regression analysis (according to the method of Lorch & Myers, 1990) that included a number of factors from three categories. The first was the auxiliary variables derived from surface form and textbase factors. These include (a) the number of syllables in a clause; (b) the serial position of the clause in the text as a whole; (c) the clause serial position within the sentence; (d) the presence or absence of new arguments in the clause; and (e) the mean frequency of the words in the clause, based on the method of Francis and Kucera (1982). The second category was the situation model variables. These include (a) changes in spatial location, (b) changes in time identified by an absolute temporal marker, (c) changes in time identified by a relative temporal marker, (d) the introduction of new discourse entities, (e) the introduction of actions without a prior causal antecedent, and (f) the introduction of a new character goal. Finally, the regression analysis included three text variables that corresponded to the different texts. The beta weights derived from these regression analyses were then used in ANOVAs to compare the age groups.

An alpha level of $p < 0.05$ was used for all statistical tests, unless indicated otherwise.

**Results**

The results of Experiment 1 showed that older and younger adults’ situation model level processing is similar. If anything, the older adults placed a greater emphasis on situation models than the younger adults did. Neither reading times nor recognition memory performance were affected by instruction. This may be either because our instructions were not strong enough or because people may be predisposed to process text of this genre, history texts, in a certain way. Because of this, none of the analyses reported here include instruction as a factor. Furthermore, although the results were not affected by vocabulary and reported prior knowledge, there was some influence of working-memory ability on performance of more propositionally related tasks, but because this effect was not replicated in Experiment 2, we do not report it here.

**Recognition**

The results of the recognition test are presented in Table 1.

A’ discrimination analysis. For older and younger adults’ discrimination of different types of information, A’ scores were calculated (see Table 1). Higher A’ scores indicate greater discrimination, and scores of .5 indicate chance performance. These data were submitted to a 2 (age) $\times$ 3 (condition) mixed ANOVA. Although the main effect of age was not significant, $F(1, 94) = 1.94, MSE = 0.028, p = .17$, there was a significant main effect of condition, $F(2, 188) = 13.13, MSE = 0.052$. In addition, the interaction was significant, $F(2, 188) = 5.73, MSE = 0.052$.

To assess the influence of age on the different representations, we performed separate analyses for each condition. Younger adults had better memory for surface form and textbase information, $F(1, 94) = 6.21, MSE = 0.040$, and $F(1, 94) = 5.19, MSE = 0.031$, respectively. This is consistent with the idea that older adults have poorer memory for the text than do younger adults. However, older adults had better memory for the situation model, $F(1, 94) = 3.95, MSE = 0.061$. This is consistent with the idea that older adults are more reliant on situation models than are younger adults. It should be noted that all of the A’ values were significantly above chance (all ps < .05), except for the surface form measure for the older adults ($t < 1$).

B’ bias analysis. Although not of primary interest, because it is conventional to present bias measures, we also calculated these as well. Positive B’ scores indicate a conservative bias, and negative B’ scores indicate a liberal bias. The B’ data were submitted to a 2 (age) $\times$ 3 (condition) mixed ANOVA. There were significant main effects of age, $F(1, 94) = 44.99, MSE = 0.442$, and condition, $F(2, 188) = 180.14, MSE = 0.180$, with older adults responding more liberally than younger adults, with the surface form responses being more liberal, and with the situation model responses being more conservative. In addition, there was a significant interaction, $F(2, 188) = 9.14, MSE = 0.180$. To explore this, we did separate analyses for each condition. Older adults were more liberal than were younger adults at the surface, textbase, and situation model levels, $F(1, 94) = 7.17, MSE = 0.175$, $F(1, 94) = 38.60, MSE = 0.325$, and $F(1, 94) = 31.05, MSE = 0.302$, respectively, with bias becoming more conservative as the deviation between the probe and the actual text increased.

**Reading Times**

The beta-weight results of the multiple-regression analyses are presented in Table 2. Each of the auxiliary variables significantly differed from zero. Thus, both older and younger adults were sensitive to these factors. The only significant age difference was...
for the number of new arguments, \( t(94) = 3.10 \), with older adults being less affected than younger adults.

Most of the situation model variables were significant, except the spatial variable for the younger adults and the intentionality variable for the older adults. It should be noted that the age difference in each case was not significant (both \( t < 1 \)). The only age difference to reach significance was for the entity dimension, \( t(94) = 2.28 \), with older adults being more affected than younger adults. None of the other age differences were significant (both \( t < 1.27 \)). It should be noted that the time variable beta weights were negative. This indicates a speed-up in reading time when temporal information was encountered. As detailed later, in the Discussion section, this is thought to be due to people ignoring the dates in the history texts, as well as any subsequent temporal information that would follow from them.

One issue of interest is the degree to which reading and recognition performance were related. The reading-time beta weights and the recognition test A' scores were submitted to correlation analyses. For Experiment 1, none of the auxiliary variables was significantly related to performance for the surface form \( (r = -0.11 \leq r \leq 0.11) \), textbase \( (-0.09 \leq r \leq 0.15) \) or situation model measures \( (-0.16 \leq r \leq 0.07) \). However, there were some significant correlations of the recognition test measures with the reading-time situation model factors.

First, there was a correlation between the absolute time factor (e.g., dates) and the surface form A' \( (r = .22) \). This is consistent with the idea that those people who were more likely to pay attention to the dates were also the ones who performed better at identifying the exact text. Closer inspection of the data revealed that this relationship was stronger for the older adults \( (r = .36) \) than for the younger adults \( (r = .20) \). This is consistent with the absolute time and surface form A' relationship described previously. None of the other variables was correlated with the textbase measure \( (-0.12 \leq r \leq 0.13) \), including absolute time \( (r = .13) \).

Finally, the situation model variables were not correlated with performance on the situation model A' \( (-0.12 \leq r \leq .19) \). However, the relationship between sensitivity to spatial information and the situation model A' was relatively strong \( (r = .19, p < .10) \). This relationship was similar for the older adults \( (r = .15) \) and the younger adults \( (r = .21) \).

**Prior Knowledge**

At the end of the experiment, people reported how much they knew about each of the topics before coming to the study. These ratings were made on a 7-point Likert-type scale, with 0 corresponding to no knowledge and 6 corresponding to complete knowledge. In general, the reported prior knowledge was very low, although younger adults reported less prior knowledge \( (M = .42) \) than did the older adults \( (M = .93), t(94) = 3.04 \).

For the A' measure, prior knowledge was not correlated with the textbase or situation model measures \( (r = -.05 \) and \( r = .10, \) respectively) but was correlated with the surface-form measure \( (r = .19) \). The more a person reported knowing, the more likely he or she was to be able to detect sentences that had actually been read. However, this difference in prior knowledge did not account for the age difference on this measure, as an analysis of covariance (ANCOVA) revealed, \( F(1, 93) = 10.93, MSE = .037 \). For the B' bias measure, prior knowledge was not correlated with the textbase and situation model measures \( (r = -.08 \) and \( r = -0.12, \) respectively, but was negatively correlated with the surface form measure \( (r = -.21) \). The more about the topics people thought they knew, the more liberal their responding was when it came to the surface

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Reading-Time Beta Weights for Experiment 1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Auxiliary variables</strong></td>
<td></td>
</tr>
<tr>
<td>Age group</td>
<td>Syllables</td>
</tr>
<tr>
<td>Young</td>
<td>.580*</td>
</tr>
<tr>
<td>Old</td>
<td>.614*</td>
</tr>
<tr>
<td><strong>Situation model variables</strong></td>
<td></td>
</tr>
<tr>
<td>Space</td>
<td>Absolute time</td>
</tr>
<tr>
<td>Young</td>
<td>.009</td>
</tr>
<tr>
<td>Old</td>
<td>.014*</td>
</tr>
<tr>
<td><strong>Text variables</strong></td>
<td></td>
</tr>
<tr>
<td>Text variables</td>
<td></td>
</tr>
<tr>
<td>Text 1</td>
<td>Text 2</td>
</tr>
<tr>
<td>Young</td>
<td>.028*</td>
</tr>
<tr>
<td>Old</td>
<td>.079*</td>
</tr>
</tbody>
</table>

\[ * p < .05 \] (marginally significant). \[ t p < .05 \]
form measure. However, this did not account for the age difference, as an ANCOVA revealed, $F(1, 93) = 4.70$, $MSE = 0.17$. In sum, with regard to recognition memory, paradoxically, prior knowledge in this study seems to have had its greatest influence on verbatim memory, not the constructive aspects of memory.

For the reading times, prior knowledge ratings were only found to be correlated with serial position ($r = .22$). The more people reported knowing about the topics in the texts, the less they were affected by the serial position of information in the text. Because there is no straightforward interpretation for this finding and because an ANCOVA revealed that this had no influence on the absence of an age difference for this factor ($F < 1$), we do not consider it further.

**Discussion**

The basic finding of Experiment 1 was that younger adults showed superior memory performance at the surface and textbase levels, whereas the older adults showed superior memory performance at the situation model level. That is, the younger adults were more adept at remembering what the text was, whereas the older adults were more adept at remembering what the text was about. This is consistent with the idea that older adults’ processing at the situation model is relatively preserved, although declines may be observed at lower levels of processing.

The other major finding of Experiment 1 was that younger and older adults showed a similar pattern with regard to their reading time data. There were only two significant age differences. The younger adults were more affected by the introduction of new arguments, whereas older adults were more affected by the introduction of a new entity. Thus, where there are significant age differences, the younger adults are more affected by changes at the surface level of the text, whereas the older adults are more affected by changes in the described situation. Although this evidence is weaker than the memory data, it is consistent with the idea that older adults’ processing at the situation model is relatively preserved, although declines may be observed at lower levels of processing.

The major results for situation model use were unaffected by reported prior knowledge (which is expected because reported prior knowledge was so low). Prior knowledge was only related to verbatim memory and serial position effects, not situation model use.

One point to note about the reading-time data were the negative beta weights for the two temporal factors. This suggests that when there was a shift in time, participants actually read faster. This is highly unusual. We suspected that this reflected people ignoring the dates in these history texts, at least in this initial pass through. As such, if they are not worrying about the specific dates at which events are occurring, they are less likely to pay attention to relative temporal markers as well. To confirm this, we looked at the processing of dates versus other temporal markers and found that these were where the largest reading-time speed-ups occurred.

The results of Experiment 1 are important in that they show both a decline in age-related performance at lower levels of processing but preserved abilities at the situation model level in the context of the same study. This supports the idea that different levels of cognition are differentially affected by aging. However, it is unclear to what extent these results are reliable and how stable they are over time. It may be that these results are unique to these materials or that older adults’ apparent preserved abilities only reflect memory when the information is needed soon after it was encountered but would falter at more substantial delays. Furthermore, there were some aspects of the results of Experiment 1 that are difficult to explain, most notably the negative beta weights in the reading-time analysis for our temporal factors (suggesting a speed up in reading when a temporal shift occurred). Experiment 2 was an attempt to generalize the findings using different materials, while at the same time trying to resolve some of the odd findings in Experiment 1 and assessing the influence of a long delay.

**Experiment 2**

One aim of Experiment 2 was to replicate Experiment 1 using different materials. There were two reasons for this. The first was to generalize the findings to a different text genre. Second, the explanation for the anomalous influence of time shifts in the reading-time data from Experiment 1 was considered (namely, the negative beta weights, indicating a reading-time speed-up). So for Experiment 2, we used narratives rather than history texts. To ensure that any observed differences were due to changes in genre, we rewrote the texts used in Experiment 1 so that they were more narrative-like. This preserved the general structure of the texts, such as the general length, causal structure, number of entities, and so on, but changed the genre. We also dropped the instructional manipulation both because it would have been odd for these texts and because we did not observe any effect of instruction in Experiment 1. In addition, no prior knowledge reports were gathered because the texts here were fictional.

Experiment 2 was also concerned with the stability of older adults’ preserved use of situation models relative to younger adults and their decline in the textbase representation. The retention of a situation model is of long-term primary importance. “When we have uttered a proposition, we are rarely able a moment afterwards to recall our exact words, though we can express it in different words easily enough. The practical upshot of a book remains with us, though we may not recall one of the sentences” (James, 1890/1950, p. 260). The retention of situation model knowledge is important because surface form and propositional knowledge are rapidly lost. To this end, we tested people at two intervals, either immediately after reading the texts or a week later. Consistent with the suggestion by James, previous research by Kintsch et al. (1990) has shown that memory for surface form and propositional information declines steadily over time, whereas memory for situation model information remains fairly stable.

**Method**

**Participants**

Seventy-two people were recruited from each of the two age groups. Half of the people from each age group were assigned to the immediate and delayed memory-testing conditions. The older adults ranged from 60 to 87 years in age ($M = 71.9$) and had 10 to 20 years of formal education ($M = 14.4$). These people were recruited from a local senior citizens’ center or were retired religious who lived on campus, and they all provided their own transportation to Notre Dame and were paid for their participation. The younger adults ranged from 18 to 25 years in age ($M = 19.5$) and had 12 to 21 years of formal education ($M = 13.5$). These people were recruited from the participant pool at Notre Dame and were given partial
class credit for their participation. The older adults had significantly more years of education than the younger adults, t(126) = 2.19. For the vocabulary test, older adults performed better (M = 33.3) than younger adults (M = 32.0), t(126) = -2.43. In contrast, for the working-memory test, younger adults performed better (M = 3.9) than did older adults (M = 2.5), t(126) = 7.86. Five additional older adults were replaced: three for failing to follow the instructions, 1 in the delay group for not returning for the second session, and 1 because of an experimenter error.

**Materials**

The four texts used were 58 to 85 sentences long and were derived from those used in Experiment 1. They were altered so that they were more narrative in style. One of the altered narratives is provided in Appendix B. In addition to the four experimental texts, there was a practice text on the topic of bullfighting that was 40 sentences long.

For the recognition test, 16 sentences were selected from each of the experimental texts. A larger set of sentences was used in Experiment 2 than was used in Experiment 1 to increase the sensitivity of our measures. Again, there were four versions of each probe sentence: verbatim, paraphrase, inference, and incorrect probes.

**Procedure**

The procedure was similar to that in Experiment 1 except for a few changes. After reading, people either proceeded immediately to the recognition test or left and were tested 1 week later. Rather than presenting the recognition test on the computer, the recognition test was administered using paper and pencil. The 16 sentences from each story were presented on a separate sheet of paper with the title of the passage at the top. The order of the sentences was randomized for each participant, and the order in which the stories were tested corresponded to the order in which they were read. The task was to indicate whether each sentence had been read before. This was done by circling either a Y or an N that appeared after each sentence. People were instructed not to go back and change their answers after they had made a selection.

**Results**

Like Experiment 1, Experiment 2 showed that the younger adults remembered the textbase information better than the older adults did. However, older and younger adults processed situation model information in a similar manner, and the older adults relied more exclusively on situation models than did the younger adults. Despite a decline in performance with the 1-week delay, older adults still retained their focus on the situation model level and showed no evidence of greater decline compared with younger adults.

**Recognition**

The recognition test data are summarized in Table 3.

A’ discrimination analysis. The A’ data were submitted to a 2 (age) X 2 (delay) X 3 (condition) mixed ANOVA. There was a significant effect of age, F(1, 124) = 34.43, MSE = 0.010, with older adults having overall lower A’ discrimination scores than younger adults. There was also a significant effect of delay, F(1, 124) = 37.58, MSE = 0.010, with discrimination declining after a 1-week delay. There was also a significant effect of condition, F(2, 248) = 43.84, MSE = 0.023, with performance increasing from surface form, to textbase, to situation model measures. Finally, there was a significant Age X Condition interaction, F(2, 248) = 13.03, MSE = 0.010. No other interactions were significant (all Fs ≤ 1.69).

To assess the influence of age on the different representations, we conducted separate 2 (age) X 2 (delay) ANOVAs for each condition. There were no significant effects for the surface form measure (all Fs ≤ 1.68). The surface form values were significantly greater than chance except for those of the older adults at delayed testing. For the textbase measure, there were significant main effects of age, F(1, 124) = 34.43, MSE = 0.024, and delay, F(1, 124) = 13.65, MSE = 0.024. This reflects the lower performance of the older adults compared with that of the younger adults and indicates that memory for the propositional content declined over time. Although performance of the younger adults was well above chance, performance of the older adults was only marginally significantly better than chance at the immediate test, t(31) = 1.75, p = .09, and not different from chance at the delayed test, t(31) = 1.12, p = .22. This result is also consistent with the idea that older adults have poorer memory for text information than do the younger adults.

Compared with performance on the surface form and textbase measures, older adults performed nominally better than younger adults on the situation model measure, although this age difference was not significant. F(1, 124) = 1.51, MSE = 0.025, p = .22. Clearly, there is no age deficit. There was a significant effect of delay, F(1, 124) = 7.03, MSE = 0.025, with performance declining over the 1-week interval. The interaction was not significant (F < 1). As should be clear from the A’ data, the older adults relied almost exclusively on their situation models to make their recognition decisions. This is evidenced by the fact that their perfor-
formance on the surface form and textbase measures were either at or near chance. In contrast, the younger adults were using a mixture of the textbase and model representations.

**Bias analysis.** The $B^*$ data were submitted to a 2 (age) $\times$ 2 (delay) $\times$ 3 (condition) mixed ANOVA. There were significant effects of age, $F(1, 124) = 11.94$, $MSE = 0.050$, and delay, $F(1, 124) = 29.99$, $MSE = 0.081$, as well as a marginally significant three-way interaction, $F(2, 248) = 2.36$, $MSE = 0.081$, $p = .10$.

To further explore these interactions, we did separate 2 (age) $\times$ 2 (delay) analyses for each measure. There were no significant effects for the surface form analysis (all $F$s $\leq 1.56$). For both the textbase and situation model analyses, there were significant effects of age, $F(1, 124) = 49.94$, $MSE = 0.163$, and delay, $F(1, 124) = 55.46$, $MSE = 0.081$, respectively, and delay, $F(1, 124) = 7.84$, $MSE = 0.163$, and $F(1, 124) = 35.73$, $MSE = 0.203$, with performance in both of these measures being more liberal for older adults and at the 1-week delay.

**Reading Times**

The reading-time data for 2 older adults were lost because of a computer error. The results of the regression analyses are presented in Table 4. For the auxiliary variables, there was a significant or marginally significant effect of each of these, except for serial position for the younger adults. There were no significant effects of age. Thus, both older and younger adults were sensitive to these factors. As can be seen in Table 4, all of the situation model variables were significant. There were also no significant age differences for these factors.

It should be noted that in an analysis using experiment as a between-subjects factor, the situation model factors had a greater influence on reading time than was the case in Experiment 1, $F(1, 218) = 55.46$, $MSE = .003$. For narrative texts, people were more concerned with the situations and events being described, whereas with history texts, people were focused more on learning the facts. The only exception to this pattern was the entity factor. A likely cause of this is the names that were used. Specifically, the names in the history texts would have been more peculiar to our readers (e.g., Thomas Knyvett) than the ones in the narrative versions (e.g., Todd Billings). This peculiarity would have lead to increased difficulty reading the names in the history texts, and hence the entity factor would have reflected a larger change in reading times.

As in Experiment 1, we examined the degree to which performance during reading was associated with performance on the memory test. As was the case with Experiment 1, none of the auxiliary variables was significantly related to performance for the surface form ($-.08 \leq r \leq .12$), textbase ($-.07 \leq r \leq .16$) or situation model measures, ($-.17 \leq r \leq .05$). The situation model factors were also not correlated with the surface form ($-.07 \leq r \leq .17$) or textbase A's ($-.16 \leq r \leq .17$). However, there was a significant correlation between the intentionality measure and the situation model A' ($r = .25$). Closer inspection revealed that this was due almost entirely to the older adults ($r = .41$), relative to the younger adults ($r = .08$). This is consistent with the idea that older adults were being driven more by their situation models in some way during comprehension and that this had a greater influence on their later memory performance.

Although the other situation model factors were not correlated with the situation model A's ($-.12 \leq r \leq .11$), the relationship between spatial shifts and the situation model memory ($r = .11$) was markedly different between the younger and older adults. The

### Table 4

<table>
<thead>
<tr>
<th>Auxiliary variables</th>
<th>Age group</th>
<th>Syllables</th>
<th>Serial position</th>
<th>Clause position</th>
<th>New arguments</th>
<th>Mean frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Young</td>
<td>-.006</td>
<td>-.018†</td>
<td>.077†</td>
<td>-.071*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Old</td>
<td>.016*</td>
<td>-.013†</td>
<td>-.021*</td>
<td>.069*</td>
<td>-.071*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Situation model variables</th>
<th>Space</th>
<th>Absolute time</th>
<th>Relative time</th>
<th>Entity</th>
<th>Causality</th>
<th>Intentionality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>.022*</td>
<td>.068*</td>
<td>.019*</td>
<td>.018</td>
<td>.056*</td>
<td>.039*</td>
</tr>
<tr>
<td>Old</td>
<td>.016*</td>
<td>.052*</td>
<td>.030*</td>
<td>.029*</td>
<td>.051*</td>
<td>.032*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Text variables</th>
<th>Text 1</th>
<th>Text 2</th>
<th>Text 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>.031*</td>
<td>-.001</td>
<td>.016</td>
</tr>
<tr>
<td>Old</td>
<td>.022*</td>
<td>.059*</td>
<td>.019</td>
</tr>
</tbody>
</table>

† $p = .08$ (marginally significant).  * $p < .05$. 

...
older adults' later memory performance was related to their sensitivity to spatial shifts \((r = 0.23)\), whereas this was not the case for the younger adults \((r = 0.01)\). Consistent with the idea of shifting processing effort from the textbase to the situation model, especially for the older adults, although the overall correlation between spatial shifts and textbase \(A'\) was not significant \((r = -0.16)\), there was a strong relationship between these factors for the older adults \((r = -0.36)\) but not for the younger adults \((r = -0.08)\).

**Discussion**

Although a significant age difference was not observed for the recognition test's situation model measure, it is clear that older adults are relying more exclusively on situation models than are younger adults. This preserved ability of the older adults remained even after a 1-week retention interval. There were no significant age differences in the reading-time measures.

The shift in the text genre had two notable effects. First, there was an increased emphasis on the situation model level, as well as a decreased emphasis on the lower levels of representation. This is consistent with the idea that when dealing with expository texts, people are thinking less about the circumstances to which these texts refer and are focusing more on their actual content (Britton, van Dusen, Glynn, & Hemphill, 1990; Graesser, 1981). In contrast with the narratives, people are more concerned with the circumstances to which the texts refer, and this was reflected in the performance data. Basically, narratives describe events, whereas expository texts describe more abstract ideas. However, both types of information are certainly involved in both.

The other effect that genre had was with regard to temporal information. In Experiment 1, temporal reading-time beta weights were negative, indicating that participants read faster when this information was encountered. A more fine-grained analysis suggested that this was due to people not paying attention to dates. When reading history texts (at least on the first pass), people are not concerned with dates. However, for narratives, reading times were positively influenced by time changes in the described situation. Thus, genre influenced how temporal information was treated.

**General Discussion**

In the context of individual experiments, with the same set of materials, for the same people, we were able to demonstrate that younger and older adults are differentially reliant on different levels of representation. The younger adults outperformed the older adults in the ability to identify information that was actually presented, whereas the older adults did at least as well as the younger adults in discriminating information concerning what the text was about. This pattern is consistent with two alternatives. First, older adults devote a larger proportion of their processing to situation models than do younger adults. Second, younger adults devote a larger proportion of their processing and memory resources to the textbase than do older adults. Both explanations are consistent with previous research that has shown that younger and older adults differ in their memory for information at lower levels of processing, such as the textbase (e.g., Cohen, 1979; Dixon et al., 1982; Meyer & Rice, 1981; Stine & Wingfield, 1988, 1990). However, they do not differ in the ability to create and use situation models (e.g., Gilinsky & Judd, 1994; Morrow et al., 1992; Radvansky et al., 1990, 1996).

**An Enhanced Younger Adult Propositional Retention Account**

Assuming a textbase deficit in older adults would be in line with the tradition of cognitive aging research. However, there may be reasons to favor a younger textbase-emphasis account. First, if the textbase is the gateway to the situation model (van Dijk & Kintsch, 1983), it is difficult to explain how a weaker textbase could lead to at least as strong a situation model as a stronger textbase. Second, younger adults in studies are often college students. College students are adapted to the task of memorizing text. As such, they may be more likely than older adults to "hang on" to the textbase because this represents what was stated in the text. Such knowledge is useful to answer multiple-choice questions. Conversely, older adults use the surface structure and textbase as more of a temporary scaffolding to construct a situation model.

A third reason is more important. In contrast to an older adult textbase-deficit explanation, a younger adult textbase-emphasis explanation accounts for both our recognition and reading-time data. During comprehension, older and younger adults appear to be equally influenced by surface features, textbase features, and aspects of the evolving situation model. Thus, the different age groups may not differentially emphasize different levels of representation during comprehension. Rather, as a younger adult textbase-emphasis explanation suggests, the younger adults retained more textbase information than did the older adults.

To assess the coherence of the younger adult textbase-emphasis idea, we implemented it in a computer simulation. We used a version of Kintsch's (1988, 1998) construction-integration model (Mross & Roberts, 1992). In this model, the three levels of representation are distinguished. During comprehension, a network is constructed consisting of nodes coding for surface elements of the text (e.g., words), textbase elements (propositions), and situational elements (e.g., tokens for protagonists and causal relations). In addition to text-derived nodes, there are also nodes from long-term memory that might get activated during comprehension.

Comprehension proceeds in cycles. During each one, part of the network is constructed and then integrated (with itself and with the previously constructed network) via a constraint-satisfaction mechanism. The network is viewed as a coherence matrix into which an activation vector is postmultiplied until the network settles. The network is considered settled when an iteration leads to a change in the network that is smaller than a preset criterion. Comprehension then proceeds to the next cycle. Typically, each processing cycle involves a clause or a sentence.

In the construction-integration model, there are several free parameters, such as the starting activation levels. Furthermore, one can "clamp" a node so that its activation remains constant. We limited our parameter selection as much as possible and used those of Kintsch et al. (1990) to model younger adults' behavior because they used the simulation to model recognition data similar to ours. These parameters were as follows: Link strengths of 5, 3, and 0 were used for nodes zero (the link of a node to itself, its self-strength), one, and two steps apart in either the surface form or the textbase. Kintsch et al. also assumed that situation-level structures are more stable and used weights of 4, 3, 2, and 1 for situational
nodes zero, one, two, and three steps apart in the situation model. Strengths of 4 were used for links between nodes at different levels (i.e., a word in the surface structure; a predicate or argument in the textbase; and a token coding for a person, object, or event in the situation model). A final assumption was that there were no direct links between the surface structure and the situation model (but see Sanford & Garrod, 1998). All of these assumptions were adopted for the younger adults.

To model the older adults, we made two assumptions: (a) surface and textbase nodes have weaker self-strengths in older adults than in younger adults and (b) more activation flows “upwards,” from the surface structure via the textbase to the situation model, than “downwards.” These assumptions reflect the idea that (a) older adults do not hang on to their surface and textbase representations as much as the younger adults do but rather (b) use the surface structure and textbase more as temporary scaffolding to construct a situation model. An important point is that no special assumptions were made about situation models relative to the younger adults. Thus, to model the older adults’ behavior, we used self-strengths of 3 and 1 for the textbase and surface structure, respectively, and retained a self-strength of 4 for the situational nodes. Furthermore, we created bidirectional links for connections between nodes from different levels. We retained the weights of 4 for upward links, from surface structure to textbase and from textbase to situation model, and assigned weights of 2 to downward links between textbase and surface structure and of 1 for downward links between situation model and textbase.

We used the following steps to model the recognition data (see Appendix C). This was based on Kintsch et al. (1990) and Kintsch (1998) in which a qualitative simulation is the aim. The purpose is to examine whether a simulation yields a pattern that is qualitatively similar to the empirical data. Typically, one case is used (i.e., one section of text with four recognition items).

1. We created two networks of the relevant section of text, one each for the younger and older adults, using the above parameters. These were the representations based on the target sentence and its predecessor, plus a clamped node for the theme: “the tulip bulb craze.”

2. We created networks for each test sentence. We used the parameters of Kintsch et al. (1990) for both age groups. Because people were viewing the test sentences, there was no decay rate.

3. Comprehension was simulated by processing the text sentences in cycles. In the first cycle, comprehension of the preceding sentence was simulated (we excluded surface nodes here for lack of relevance to our current purposes). This resulted in an integrated network consisting of textbase and situational nodes and their links.

4. In the next cycle, target sentence comprehension was simulated and integrated with the representation based on the prior sentence. We assumed that readers routinely infer causal antecedents (Graesser, Singer, & Trabasso, 1994). Consequently, a proposition and a situational node were included, representing the causal antecedent: “tulip bulbs look like onions.”

5. The integrated representation of this portion of the text was moved to long-term memory (LTM).

6. Each test sentence was integrated in separate simulations with the LTM network. The test sentence network was linked with the LTM structure so that each test sentence node was linked to its counterpart—when present—in LTM, using the same parameters that were used for LTM representations. Thus, all of the nodes of the verbatim sentence had a counterpart in LTM. Some of the surface elements and all of the textbase and situational nodes of the paraphrase had a counterpart in LTM. None of the surface nodes, some of the textbase nodes, and all of the situational nodes of the inference had a counterpart in LTM. Finally, only one textbase and one situational node of the incorrect item had a counterpart in LTM.

7. Finally, we summed the activation levels for all of the test-sentence nodes and divided by the total amount of activation in the network (test sentence plus activated part of LTM) to control for normalizing (Kintsch et al., 1990).

Figure 1 conveys one of the networks used to represent the older adults’ representation of a segment of the tulip bulb craze text plus a test sentence. The network consists of three parts, which are vertically organized. The top part represents the sentence preceding the target. The middle represents the target sentence. Finally, the bottom represents the test sentence, in this case an inference sentence. The squares (labeled V for verbatim) are nodes representing the surface structure. The circles (labeled P for propositional) are textbase nodes. The triangles (labeled M) are situation model nodes. Finally, nodes labeled TP (for test proposition) and TM (for test model) are textbase and situation model nodes from the test sentence. No verbatim nodes were included for the preceding and test sentences. The reason was that they would not be connected to the target sentence and therefore were largely irrelevant.

The links’ thickness represents the connection strengths: The thicker the line, the stronger the link. The links from different levels are bidirectional. For example, a link emanating from a textbase node to a situation model node is thicker on the textbase side than on the model side, indicating that more activation flows from the textbase node to the model node than vice versa.

Two nodes need to be discussed in more detail. Node M1 represents the topic of the text, the tulip bulb craze, which we assumed was part of the representation in view of empirical evidence (e.g., Lorch, 1993) but which is not essential to the simulation. This node was clamped so that its activation level did not vary across cycles. This reflects the special status of topic nodes in the memory representation (Guindon & Kintsch, 1984). The second node, which is of central importance to the simulation, is M5, which represents the inference that tulip bulbs look like onions. This, of course, is also part of the test sentence represented here (Node TM2). Consequently, there is a link from TM2 to its counterpart in the memory representation M5.

The corresponding network for the younger adults was identical in structure, but the connection weights were different, as discussed earlier. The networks for the verbatim and paraphrase test sentences also included surface nodes for the test sentence, which were connected to their counterparts in LTM. Figure 2 presents the outcome for younger and older adults. On the left y-axis are the simulation’s activation levels (multiplied by 1,000). On the right y-axis are the probabilities of a “yes” response in Experiment 1. There is a nice qualitative fit between the simulation and empirical data. Both groups are similar in their responses to verbatim items, but the older adults are more likely to respond “yes” to paraphrases and particularly more likely to respond “yes” to correct inferences and also more likely to respond “yes” to incorrect inferences.
Figure 1. One of the networks used in the simulations to represent the older adults' memory representation of a segment of the "Tulip Bulb Craze" text plus a test sentence. The network consists of three parts, which are vertically organized. The top part is the representation of the sentence preceding the target sentence. The middle part represents the target sentences. Finally, the bottom part represents the test sentence, in this case the inference sentence. The squares (labeled V for verbatim) are nodes representing the surface structure. The circles (labeled P for propositional) are textbase nodes. The triangles (labeled M) are situation model nodes. Finally, nodes labeled TP (for test proposition) and TM (for test model) are textbase and situation model nodes from the test sentence. The thickness of the links between the nodes represents the connection strengths: The thicker the line, the stronger the link. The links between the nodes from different levels are bidirectional. For example, a link emanating from a textbase node to a situation model node is thicker on the textbase side than on the model side, indicating that more activation flows from the textbase node to the model node than vice versa. The corresponding network for the younger adults was identical in structure, but the connection weights were different. The networks for the verbatim test sentence and paraphrase also included surface nodes for the test sentence, which were connected to their counterparts in the long-term memory representation.

This simulation provides computational support for a younger textbase-emphasis account. By assuming that older adults (a) have weaker self-strengths for surface and textbase nodes than do younger adults and (b) have more activation flowing from lower levels of representation to higher levels than vice versa, whereas younger adults show no directionality in this respect, we were able to obtain a good qualitative fit with the empirical data of Experiment 1.

Of course, this does not prove that this possibility is the correct one. Other possibilities may also be operating, all of which could be true. In the next section, we present three other possible accounts of our findings.

Other Alternatives

There are three other explanations for older adults' preserved abilities to create and use situation models. One is that older adults are more selective in the information they process during comprehension. From this perspective, comprehension is a skill, with experts being better able to select that information that is important
for understanding (Chase & Ericsson, 1982). The goal of comprehension is to create a situation model. Older adults are more practiced at comprehension and thus have more expertise than younger adults at achieving this. This allows them to identify those portions of a text pertinent to model construction. For example, in Experiment 1, the reading times of older adults tended to be more sensitive to information that was relevant to model construction (e.g., the introduction of new entities).

A second idea is that memory and comprehension reflect a greater reliance on schematic knowledge. The research on older adults' schema use is mixed. Although it is clear that younger and older adults' schemas have similar content and structure (Hess, Vandermaas, Donley, & Snyder, 1987; Light & Anderson, 1983), the extent to which they are dependent on them is not. Some researchers have suggested that older adults are as dependent on schemas as younger adults are (Arbuckle, Cooney, Milne, & Melchoir, 1994; Arbuckle, Vanderleck, Harsany, & Lapidus, 1990; Hess & Flannagan, 1992; Hess et al., 1987; Light & Anderson, 1983; Zelinski & Miura, 1988). In contrast, others have suggested that older adults may place a greater emphasis on them (Arbuckle et al., 1994; Labouvie-Vief & Schell, 1982; Smith, Rebok, Smith, Hall, & Alvin, 1983), including during comprehension (Miller & Stine-Morrow, 1998). This notion of increased dependence on schematic knowledge is consistent with Reder, Wible, and Martin's (1986) idea that older adults are more likely to use plausibility strategies during retrieval in contrast to more exact retrieval strategies.

A third idea is that older adults generate, or at least maintain, a wider range of inferences during comprehension (Hamm & Hasher, 1992). The creation of a situation model is essentially an inference-making process in which the given information and general world knowledge is used to construct an understanding of the described situation. Studies of story telling have found that older adults are generally more verbose than are younger adults (e.g., Arbuckle & Gold, 1993) and that this speech is more likely to be off the main topic (although not always). However, older adults' stories are also rated as more informative, interesting, and of a higher quality (James, Burke, Austin, & Hulme, 1998). Thus, older adults are bringing more information to their understanding of an event (increased inferences). This is consistent with the idea that they are placing a greater emphasis on developing situation models.

Also consistent with this idea that older adults are creating more inferences is work by von Hipple, Silver, and Lynch (2000). In this study, measures of prejudice, impression management, and inhibition were obtained. Older and younger adults then read a narrative about a student; the narrative was biased toward particular racial stereotypes. People were explicitly told to form an impression of the student and to not allow social category information to influence their responses. The results showed that older adults were more prejudiced than were the younger adults, inferring more information based on the student's racial background. This occurred despite the fact that the older adults had a greater desire to avoid being prejudiced as reflected in their higher impression management scores. A more detailed analysis revealed that performance was related to inhibitory ability but not impression management or prejudice.

**Expanded Horizons**

The current pattern of results is consistent with some research reported by Koutstaal and Schacter (1997) using a variant of the Deese–Roediger–McDermott false-memory paradigm (Deese, 1959; Read, 1996; Roediger & McDermott, 1995). In Koutstaal and Schacter's study, participants were presented with a series of line drawings. Three days later, they were given a recognition test composed of studied, categorically related, and unrelated items. Performance on these different items were then used to calculate $A'$ scores, as was done in the current study.
As in the current study, for item-specific memory (studied items vs. categorically related items and studied items vs. unrelated items), the younger adults outperformed the older adults. However, for gist memory (categorically related items vs. unrelated items), the older adults outperformed the younger adults. This parallels our findings. Also as in the current study, the B' measure showed that older adults were more liberal responders than were the younger adults.

Koutstaal and Schacter (1997) interpreted their results as showing that older adults are more susceptible to categorical similarity. Although their basic pattern is similar to ours, this explanation is not a satisfactory account of our data. Specifically, our inference probes, unlike their related items, were not categorically related to the studied items. Instead, the relationship involved more causal relations, goals, and so forth, not feature overlap. Likewise, it is unlikely that a situation model explanation can be applied to Koutstaal and Schacter’s data because their materials were not situation specific—pictures of objects in the absence of any context.

However, at a more global level of explanation, both these accounts are consistent with fuzzy-trace theories of memory (Brainerd, Reyna, & Mojaridin, 1999). In this view, information is stored in at least two levels of representation: a detailed, item-specific trace and a coarse-grained, gist-based representation. Memory performance reflects the combined influences of these representations. In our study, the item-specific information was the idea units presented in the text (the textbase), and the gist-based representation was the category to which those objects belonged. In both cases, older adults exhibited a reduced emphasis on the item-specific representation and a greater emphasis on the gist-based representation. It is unclear whether this shift in processing emphasis is due to increased expertise or to deficits at lower levels of processing. Moreover, it is unclear why this shift occurs.

Conclusions

The results of our study clearly demonstrate that there are preserved abilities in older adults’ situation model processing even though there are declines in processing at lower levels, such as the retention of verbatim or propositional information. This occurs despite declines in older adults’ cognitive functioning more generally, such as declining working-memory capacity. There is even the possibility that a nominal age-related decline, namely decreased inhibitory ability, may be partially responsible for the preserved situation model processing. More research is needed to attack these questions and bring some insight into the nature of the preserved abilities of the older adults and into how these preserved abilities might be used to compensate for the consequences of declines that accompany aging in other areas of cognitive functioning.

References


Appendix A

Sample Test Passage: The Gunpowder Plot

The gunpowder plot in 1605 is an important episode in the history of England. It stemmed from a set of long-standing grievances. English Roman Catholics had been subjected to severe penal laws. These laws had been imposed during the reign of Queen Elizabeth I, and when her successor James I came to the throne in 1603, some relief had been hoped for and, indeed, promised. When it was not forthcoming, Robert Catesby, a member of the Catholic gentry, proposed to right the wrongs by blowing up the Parliament house. The goal of the plotters was to blow up Parliament and King James I, his queen, and his oldest son on Nov. 5, 1605. This would provide an opportunity for the English Catholics to take over the country. Catesby enlisted his wealthy cousin Thomas Winter in his scheme and sent him to Flanders in April 1604 to hire the services of an English Roman Catholic soldier of fortune, Guy Fawkes, to carry out the plot when Parliament met at the beginning of the following year. On May 24 a fellow conspirator, Thomas Percy, hired a tenant adjoining the Parliament house in furtherance of a plan to dig a tunnel through the dividing wall. A second house, for storing gunpowder, was rented in Lambeth. In March 1605, however, the conspirators were also able to rent a coal cellar that actually ran under the Parliament house. Fawkes carried 20 barrels of powder into the cellar and covered them with firewood until the meeting of Parliament, when the deed was to be done. The conspirators then separated until the meeting of Parliament. The meeting of Parliament, meanwhile, had been postponed to November. In the interim the need for broader support persuaded Catesby to include more conspirators. So, he brought in other conspirators, including another cousin, Francis Tresham. Altogether, 13 Catholics were directly engaged in the plot. A difficulty was that the explosion might kill friendly Catholic members of the House of Lords. Tresham was particularly anxious to warn his brother-in-law, Lord Mon-teagle. On October 26, Monteagle showed a letter to the King’s principle minister, Lord Salisbury, who in turn showed it to the King. It was decided to search the Parliament house and the adjoining buildings. The search was conducted on November 4, first by Lord Suffolk, who actually encountered Fawkes in the cellar, and saw the piles of firewood, and that night by Sir Thomas Knyvett, a Westminster magistrate, who discovered the barrels of gunpowder and arrested Fawkes, who, under torture, confessed and re­vealed the names of the conspirators. The other conspirators fled from London but were rounded up in Staffordshire. Catesby and Percy were killed, and all of the others that were involved were tried and executed on January 31, 1606. The plot bitterly intensified Protestant suspicions of Catholics. It led to the rigorous enforcement of the recusancy law, which fined those who refused to attend Anglican services. In January 1606 Parliament established November 5 as a day of public thanksgiving. The day, known as Guy Fawkes Day, is still celebrated with bonfires, fire­works, and the carrying of “guys” through the streets. There are many mysteries about the Gunpowder Plot that have never been satisfactorily explained. It is not clear why an intelligent man like Catesby thought that such a scheme would work or why he imagined that if it did work, a small group of Catholics could seize the reigns of government. Nor is it clear why the King and Lord Salisbury immediately interpreted Monteagle’s letter to mean that a gunpowder plot was intended. Also it is odd that although the letter was received on October 26, the search of the cellars was not carried out until November 4. Nevertheless, the detailed confessions that have survived, including that of Guy Fawkes, make it difficult to believe, as has been argued, that the whole story was invented by Lord Salisbury to strengthen his position in the government of James I.

Appendix B

Sample Test Passage: The Farmers’ Rebellion

Last summer’s farmer rebellion was an important event in our town of Pitman. It stemmed from a set of long-standing grievances. Many farmers had been subjected to severe environmental laws. These laws had been implemented during the term of mayor Judy Copeland, and when her successor James I came to the throne, some relief had been hoped for and, indeed, promised. When it was not forthcoming, Robert Catesby, a member of the Catholic gentry, proposed to right the wrongs by blowing up the Parliament house. The goal of the plotters was to blow up Parliament and King James I, his queen, and his oldest son on Nov. 5, 1605. This would provide an opportunity for the English Catholics to take over the country. Catesby enlisted his wealthy cousin Thomas Winter in his scheme and sent him to Flanders in April 1604 to hire the services of an English Roman Catholic soldier of fortune, Guy Fawkes, to carry out the plot when Parliament met at the beginning of the following year. On May 24 a fellow conspirator, Thomas Percy, hired a tenant adjoining the Parliament house in furtherance of a plan to dig a tunnel through the dividing wall. A second house, for storing gunpowder, was rented in Lambeth. In March 1605, however, the conspirators were also able to rent a coal cellar that actually ran under the Parliament house. Fawkes carried 20 barrels of powder into the cellar and covered them with firewood until the meeting of Parliament, when the deed was to be done. The conspirators then separated until the meeting of Parliament. The meeting of Parliament, meanwhile, had been postponed to November. In the interim the need for broader support persuaded Catesby to include more conspirators. So, he brought in other conspirators, including another cousin, Francis Tresham. Altogether, 13 Catholics were directly engaged in the plot. A difficulty was that the explosion might kill friendly Catholic members of the House of Lords. Tresham was particularly anxious to warn his brother-in-law, Lord Mon-teagle. On October 26, Monteagle showed a letter to the King’s principle minister, Lord Salisbury, who in turn showed it to the King. It was decided to search the Parliament house and the adjoining buildings. The search was conducted on November 4, first by Lord Suffolk, who actually encountered Fawkes in the cellar, and saw the piles of firewood, and that night by Sir Thomas Knyvett, a Westminster magistrate, who discovered the barrels of gunpowder and arrested Fawkes, who, under torture, confessed and re­vealed the names of the conspirators. The other conspirators fled from London but were rounded up in Staffordshire. Catesby and Percy were killed, and all of the others that were involved were tried and executed on January 31, 1606. The plot bitterly intensified Protestant suspicions of Catholics. It led to the rigorous enforcement of the recusant law.
which fined those who refused to adopt government policies. This summer, the city council established August 5th as a public holiday. That day, known as Field’s Day, was celebrated with bonfires, fireworks, and the carrying of “pigs” through the streets. There are many mysteries about the farmers’ rebellion that have never been satisfactorily explained. It is not clear why an intelligent pig farmer like Collins thought that such a scheme would work, or why he imagined that if it did work, a small group of farmers could seize the reigns of the government. Nor is it clear why the mayor and Steve Flett, his lawyer, immediately interpreted Thorn’s letter to mean that an explosion was intended. Also, it is odd that although the letter was received on July 26th, the search of the cellar was not carried out until August 4th. Nevertheless, the detailed confessions that have been made public, including that of George Fields, make it difficult to believe, as has been argued, that the whole story was invented by Steve Flett to strengthen his position in the government of Pitman.

Appendix C
Sentence-Processing Examples

Sentence Placement Example

Previous sentence: “People who had been away from Holland and then returned during the craze sometimes made mistakes.”

Target sentence: “It is said that a sailor mistook a tulip bulb worth several thousand florins for an onion.”

Paraphrase: “It is said that a sailor accidentally thought that a tulip bulb worth several thousand florins was an onion.”

Correct inference: “Tulip bulbs often resemble other bulbs, such as onions and garlic.”

Incorrect inference: “People often stored the bulbs in secure places.”

Test Sentences

Verbatim: “It is said that a sailor mistook a tulip bulb worth several thousand florins for an onion.”

Received October 4, 1999
Revision received June 28, 2000
Accepted June 29, 2000