Journal of Gerontology: PSYCHOLOGICAL SCIENCES
1997, Vol. 52B, No. 4, P187–P195

Adult Age-Group Differences in Recall for the Literal and Interpretive Meanings of Narrative Text

Cynthia Adams,[1] Malcolm C. Smith,[2] Linda Nyquist,[3] and Marion Perlmutter[4]

[1]Center for the Study of Women in Society, University of Oregon.
[2]Department of Marketing, University of Manitoba.
[3]Institute of Gerontology and [4]Department of Psychology, The University of Michigan.

This study examined age differences in recall for the literal and interpretive meanings of narrative text. Following presentation of one of two stories rich in both literal and interpretive content, younger (mean age = 19.2 years) and older (mean age = 72.2 years) adults were asked to retell and to interpret the story. Response task order was counterbalanced across participants. When asked to retell a story as close to the original as possible, the younger adults recalled more of the literal propositional content than did the older adults in the retell-first, although not in the interpret-first, condition. In addition, both older and younger adults recalled more of the main ideas (gist) relative to the details. When asked to interpret the same story, more older than younger adults produced deep and synthetic representations of the story's interpretive meanings. In addition, there was a clear preference among the older age group for deep-synthetic responding. Although more younger than older adults produced analytic interpretations, within the younger group there was no clear preference for either an analytic or a deep-synthetic style.

IN the past few years, there has been an emergent interest in the question of qualitative age differences in the way text information is processed. It remains to be established, however, what, if any, those qualitative differences might be (Stine & Wingfield, 1990). As many studies have shown, we know that as people grow older, detailed recall drops out in favor of recalling the gist of a text (Hultsch & Dixon, 1984; Meyer & Rice, 1989). This age-related change is often regarded as a compensatory mechanism for declines in information-processing efficiency.

Our own work suggests that this gist-detail trade-off may be only part of the picture. For example, in studies on narrative recall we have observed that older adults differ from younger adults in two distinct ways. First, when asked to freely recall a story, older adults often spontaneously reconstruct portions of the literal version of the text into their deeper interpretive meanings. Second, older adults often produce briefer versions of the story — not simply reproductions of the higher order propositions (the criterion for gist), but rather highly integrative and succinct representations of a story's essential meaning (Adams, 1991; Adams, Labouvie-Vief, Hobart, & Dorosz, 1990; see also Labouvie-Vief & Schell, 1982). Much as experienced chess players make sense of a chessboard by processing patterns of moves rather than individual pieces (Dreyfus & Dreyfus, 1986), experienced text-processors (i.e., older adults) seem to be making sense of text information by processing patterns of meaning rather than individual propositions. On the basis of these observations, we and other authors (e.g., Labouvie-Vief, 1990) suggest that detailed recall might drop out because the deeper psychological and symbolic meanings of narrative grow more salient with age (see also Carstensen & Turk-Charles, 1994), and because older adults process these elements in a succinct and integrative way.

As we have argued elsewhere (see Adams, 1991; Adams et al., 1990), without research designs and methods capable of tapping such an interpretive and integrative processing style, older adults' processing abilities are probably underestimated. Research on memory for text in adulthood focuses primarily on the comprehension and recall of the propositional representation of meaning (for reviews, see Hultsch & Dixon, 1984, 1990; Meyer & Rice, 1989). Because propositional representations of meaning are closely tied to the surface-level text (see also Holyoak, 1982; Johnson-Laird, 1983), recall performance is most often evaluated in terms of an individual's ability to reproduce a text's literal meaning. Consequently, the most credit is given for the most propositional details recalled.

While this exclusive focus on memory for the literal and detailed meaning of a text may not present significant problems for the study of expository text, it may be problematic for the study of the narrative form. This is because most stories contain a deeper psychological and often symbolic "landscape" of meaning (see Bruner, 1986) — a layer of meaning that is not explicitly represented in the propositional text per se, but is interpretable from that text. Narrative is thus more open to the interpretive and integrative reconstructions of the reader than is exposition (Adams, 1991). For a truly in-depth comprehension of a story to take place, narrative processing should require the reader to comprehend both the propositional and deeper-level meanings (Dyer, 1983; Holyoak, 1982). Yet, recall of such deeper meanings has been largely ignored in the text-memory and aging literature.

It is in studies on narrative recall, then, that the exclusive

focus on the literal propositional text may put the older information-processor at a disadvantage. As noted previously, older adults seem to be more interested in, and thus more attentive to, the deeper meaning or significance of an event than to the individual details comprising the event (Adams, 1991; Jepson & Labouvie-Vief, 1992; Labouvie-Vief, 1990). Moreover, older adults are less likely than younger adults to take a detailed and analytic approach to the text-processing task (Birren, 1969; Boswell, 1979; Kramer & Kahlbaugh, 1989; Labouvie-Vief & Schell, 1982), favoring instead a more integrative or synthetic text-processing style.

Hence, when reading or listening to a story, the older information-processor either may selectively ignore and thus not encode as many of the propositional details as do younger adults or may experience the rapid decay of surface-level codes as these elements become transformed and integrated into deeper meaning codes (Labouvie-Vief & Schell, 1982). In either case, when asked to retell a story as close to the original text as possible, older adults would be expected to produce fewer surface-level propositions than younger adults. And, in most studies of narrative recall in adulthood, such age-related differences are the norm (see Hartley, 1989; Zelinski & Gilewski, 1988). Such findings are most often interpreted as evidence for processing deficits associated with aging (Cohen, 1988; Zelinski & Gilewski, 1988). When asked to interpret the same story, however, an expanded picture is likely to emerge: Older adults should exhibit an ability to comprehend and recall the story's deeper interpretive meanings.

To better understand the text-processing abilities of older adults, it seems necessary to examine age-group differences for both literal propositional and deeper interpretive levels of meaning in narrative text. The present study was designed to do so. Specifically, younger and older adults were asked to both retell and interpret one of two stories, each containing deep interpretive meanings. The retelling responses were used to assess recall of the story's literal meaning. Because propositional representations of text are closely tied to the literal surface-level of a text, recall for the literal meaning of a story was evaluated in terms of the amount of propositional content present in the retellings. Both total recall and recall for gist relative to details were analyzed.

The interpretation responses were used to assess recall for a story's interpretive meanings. We assumed that both younger and older adults are capable of interpreting stories; however, we expected the quality of the interpretations to vary by age. If there is an age-related progression toward deep, more integrative processing, and if older adults are more adept at interpreting stories, we reasoned that they, as experts in the interpretive domain, would be more likely than younger adults to produce deep and synthetic representations of a story's interpretive meanings. Behaving more like novices, younger adults were expected to take a more analytic and perhaps more literal approach, stringing together, one by one, the interpretive elements of a story rather than achieving a patterned whole or synthesis of those elements. Thus, the quality of the interpretations was evaluated by two criteria: (a) the extent to which a response reflected the deep meaning(s) of a story (depth criterion), and (b) the extent to which a response achieved an integrated synthesis of the interpretive elements in the story (synthesis criterion).

We predicted the following pattern of age-related differences: (a) for recall of the literal meaning of a story, younger adults were expected to produce more propositional content than older adults; (b) for recall of the gist of a story, both older and younger adults were expected to show better recall of the main ideas relative to the details; and (c) for recall of the interpretive meanings of a story, more older than younger adults were expected to produce deep and synthetic interpretations.

Method

Participants

Twenty-four younger adults ranging in age from 18 to 21 years (mean age = 19.2 years) and 24 older adults ranging in age from 64 to 77 years (mean age = 72.2 years) volunteered to participate in the study. Half of the participants in each age group were women and half were men. The two age samples were recruited from similar populations: The younger adults were undergraduates at the University of Michigan who were recruited from a number of psychology classes, and the older adults were University of Michigan alumni who responded to a mailing requesting volunteers for cognitive aging research. Because one young male participant reported that he had been exposed to the study hypothesis in one of his classes, he was dropped from the study.

Table 1 compares the two age groups on several relevant characteristics. Although the age groups differed in years of education completed ($t = 10.47$, $p < .001$), they did not differ in either socioeconomic status or verbal ability. Because verbal ability is a known correlate of text-memory performance, we felt confident in the general equivalence of the two groups on a relevant cognitive variable.

Because self-reported health is correlated with cognitive functioning in later life (Perlmutter & Nyquist, 1990), participants were asked to complete the 13 items of the Cornell

Table 1. Descriptive Data for Participants by Age Group

	Older Adults ($n = 24$)		Younger Adults ($n = 23$)	
Variable	*M*	*SD*	*M*	*SD*
Age	72.20	2.93	19.20	0.85
Education (in years)	17.79	1.72	13.39	1.08
Social status score[a]	60.43	4.67	57.13	8.27
Vocabulary score[b]	32.58	4.13	30.95	5.67
Cornell Medical Index	1.21	0.88	0.26	0.45
Health rating[c]	100.00		100.00	
Vision rating[c]	92.00		100.00	
Hearing rating[c]	88.00		100.00	

[a]Based on Hollingshead's (1975) Four-Factor Index. Values can range from a low of 6 to a high of 66.

[b]Based on Jastak and Jastak's (1964) 20-item written form of the Wechsler Adult Intelligence Scale Vocabulary subtest. A total of 40 points is possible.

[c]Percentage of participants with "good" or "excellent" self-ratings.

Medical Index (a self-report symptom checklist; Brodman, Erdman, & Wolff, 1949) dealing with cardiovascular illness. In addition, participants rated their present eyesight, hearing, and overall health on a scale from 1 (excellent) to 4 (poor). As shown in Table 1, the older adults, on average, reported only about one health symptom, and the majority of participants in each age group rated themselves as good or excellent on the health scales.

Stimulus Texts

Two narratives were used as the stimulus materials in this study. Each was a modified version of a Sufi teaching tale (Shah, 1967). This type of narrative was selected because it contains at least two levels of meaning for potential analysis and encoding. At one level a Sufi tale is a relatively simple action-event sequence represented in the text's propositional content. At another level a Sufi tale is rich in psychological content represented in several levels of symbolic meanings (Shah, 1967). Such narratives thus maximize the potential salience of both the literal and interpretive meanings in text.

One story was about a personified stream (Stream story) who wanted to cross a desert in order to continue its journey across the countryside. The other story was about a monk (Monk story) who wanted to show a second monk the correct way to repeat a chant. The Stream story contained 467 words, and the Monk story contained 421 words. Both texts fell within the 7th-grade readability range as assessed by the Fry (1968) method. Using criteria developed by Turner and Greene (1977; based on Kintsch, 1974), a propositional analysis was carried out on each text. The text base of the Stream and Monk stories contained 194 and 188 propositions, respectively.

Procedure

Participants were tested individually in a university laboratory. They were first interviewed to obtain demographic information and to establish rapport between the participant and researcher. They then completed the health questions. Within each age and sex grouping, participants were randomly assigned to either the Stream or the Monk story condition and were read the following instructions:

> Now I would like you to read a short story. Following the story, you will be asked to remember the story in two ways. I will be asking you to both retell the story itself and to say what you think the story means; that is, to interpret the story. So, please read the story carefully so that you can remember it. You may take as much time as you need to read the story. However, please read it through only once. Just let me know when you are finished.

Reading time was recorded with a handheld digital stopwatch.

Following presentation, the story was removed and a 10-min filler task was administered to prevent participants from thinking about the story, thus avoiding short-term memory rehearsal. The Physiognomic Sensitivity card-sorting task (Kogan, Connor, Gross, & Fava, 1980) was used for this purpose. If the task was completed before the required 10 min lapsed, the participant was interviewed about the strategies he or she used while doing the card-sorting task until the full 10-min time period elapsed.

Immediately following the filler task, half of the participants (within each Age × Sex × Story grouping) were asked to retell as much of the story as they could remember and as close to the original text as possible. On completion of the retell task, they were asked to interpret or give the meaning of the story. The remaining half of the participants completed the two response tasks in the reverse order, thus creating a counterbalanced design. Responding was oral and was audiotape-recorded. Participants were advised that there were no time limits on these two tasks.

On completion of the two response tasks, participants were asked to rate the story from 1 (not at all) to 5 (very) on each of the following scales: (a) the extent to which they found the meaning of the story difficult to understand, (b) the extent to which they were familiar with this type of story, (c) the extent to which they liked the story, and (d) the extent to which they identified with a character or characters in the story. They were then asked to complete a vocabulary measure (see Table 1).

Response Scoring

In preparation for coding, each retelling and interpretation response was transcribed from the audiotape and assigned an arbitrary code so that coders would be unaware of the age and sex of the participant.

Retelling responses. — To assess recall for a story's literal text, retelling responses were checked against their respective text base for the presence of each proposition. Using a lenient criterion (see Turner & Greene, 1977), one coder scored the entire set of retellings. To check the reliability of the propositional analysis, a second coder scored a randomly drawn sample of 25% of the responses (six from each story). There was 87.34% agreement between the two coders.

To assess recall of the gist relative to the details in a story, the propositions in each text were rated for levels of importance by two of the authors. First-level propositions were those that stated the main ideas of the story. Second-level propositions were supporting statements of the main ideas. Third-level propositions were supporting details, and fourth-level propositions were inessential details. Differences in the ratings were resolved in conference to provide a final set for each story. For the Stream story, the resulting number of propositions for Levels 1 through 4 was 46, 55, 44, and 49, respectively; for the Monk story, the number of propositions at each level was 42, 53, 50, and 43.

Interpretation responses. — To assess the deep and synthetic qualities of interpretive recall, we developed two 5-point rating scales, a depth scale and a synthesis scale, with 1 designating the low end of a scale and 5 designating the high end. The depth scale was designed to assess the extent to which an interpretation represented a deep covert meaning of a story. A low rating indicated a relatively shallow or literal interpretation, one that remained tied to a story's propositional text base. A high rating indicated a

deep or symbolic interpretation, one that went beyond the overt content and structure of the propositional text to express the deep symbolic themes in the story. Following are examples of a shallow and deep interpretation of the Stream story:

Shallow interpretation (low on depth scale):

> It seems like it's a story about cycles. It's a story about a stream that doubted that he could ever do something different and still be the same, still be a stream. And if there was nothing that could convince him that no matter what he did, if he just stayed that way, if he didn't try to change, to get by, then he wouldn't be able to be something more than a stream. And the stream realized that just because of a small change doesn't mean that he won't be a stream anymore. He'll still be a stream, but he'll still move and change and be better than just a swamp.

Deep interpretation (high on depth scale):

> I think I see the story on three different levels. The first interpretation is that you have to be willing to be interdependent. Be willing to work with people, and not stand alone. 'Cause if you stand alone, then you are alone and you don't get, you don't prosper as much, and you don't feel the good feelings that you get out of being with people. The second thing is that you have to be willing to accept change. You can't be rigid and set into your own created set of values and principles. And your knowledge, you have to be willing to change with the changes the circumstances require. And then the third is sort of a synthesis of the two. That you make change most easily and effectively if you work with people. That you don't need to make change on your own.

The synthesis scale was designed to assess the extent to which an interpretation represented a complete and integrated whole. A low rating indicated an analytic response, or one in which the focus of the interpretation was on delineation, explanation, and/or illustration of the individual concepts or symbols in a story. A high rating indicated a more integrated response in which the symbolic elements of a story were unified into a succinct and complete whole. Synthetic responses were conceptually dense, abstract, and holistic. Following are examples of an analytic and a synthetic interpretation of the Stream story:

Analytic interpretation (low on synthesis scale):

> I would interpret the story to be a representation of how a person would deal with a barrier or problem put in front of him. The barrier being the desert and the person's life being the stream. And the wind would be religion or faith in something as a means of overcoming that barrier to achieve what your objective is.

Synthetic interpretation (high on synthesis scale):

> The story is kind of a parable. It is intended to set the tone for awakening people to aspects of their personality which they had not been aware of in the past. The parable indicates that maybe an outsider can make suggestions about a person's own character which will awaken that person to some strengths which he didn't know he had and can suggest ways of using those parts of his character in conjunction with other things in his environment to overcome difficulties. And in the end, the person who is having difficulty becomes aware, number one, of additional strengths and, number two, of the fact that he can use his environment to carry himself forward.

Note that although the above analytic example is relatively short, by its very nature the analytic approach to the task was expected to produce lengthier interpretations than those produced by the more integrated synthetic approach. Length was therefore expected to be negatively correlated with the ratings on the synthesis scale. Length was not, however, a necessary condition for an analytic response and was therefore not a defining feature for the coders to use when rating interpretations on the synthesis scale.

Nine coders were trained to score the interpretation responses. Training procedures were adapted from those developed by White (1985) for use in the "holistic" scoring of answers to essay questions used in tests constructed by the Educational Testing Service. All coders were brought together at the same time and place for training as well as for coding. At the first session, coders read through the entire set of protocols to become familiar with the data. At the second session, a coding criteria guide was introduced, and coders were asked to read through the set of protocols a second time, keeping the scoring criteria in mind while reading. The coding guide gave a general definition of each criterion as well as a description of the specific features that defined end- and midpoints on the two scales. At the third session, sample protocols were distributed, independently coded, and discussed. This series of controlled readings served two purposes: (a) to ensure that all coders were familiar with, and had deeply processed, the criteria prior to coding, and (b) to facilitate a sense of community among the coders which, according to White, is essential to the success of this type of coding.

The final coding was also conducted in group sessions, with a separate session designated for each story. Within each session, three or four coders were assigned to each of the two rating scales. Once assigned, each coder worked independently until he or she completed the entire set of response protocols. These group coding sessions not only eliminated possible extraneous factors in the scoring procedure but also served to further enhance the sense of community among the coders.

Interrater reliabilities were assessed using Cronbach's alpha. The alpha coefficients for the Stream story were .88 for depth and .84 for synthesis. For the Monk story, alphas were .77 for depth and .89 for synthesis.

Results

Preliminary Analyses

Three sets of preliminary analyses were conducted. First, to determine whether the two age groups differed on any of the story rating variables, a 2 (age group) × 2 (story) × 2 (response task order) multivariate analysis of variance (MANOVA) was conducted with the four rating scale scores as the dependent variables. Response task order was included in the analysis because participants were asked to rate the stories following completion of the two response tasks. Only the multivariate interaction of Age Group × Response Task Order was significant, Wilks' lambda = .67, $p < .01$,

with univariate tests significant for the difficulty and familiarity ratings only, $Fs(1,38) = 5.82$ and 6.48, respectively, $ps < .05$. A response task order effect was found within the older, but not the younger, age group. Specifically, the older adults rated the stories as more difficult in the retell-first condition ($M = 1.91$, $SD = .94$) than in the interpret-first condition ($M = 1.17$, $SD = .39$), simple effects test, $F(1,38) = 4.96$, $p < .05$. They also rated the stories as less familiar when they were required to retell the story first ($M = 2.83$, $SD = 1.26$) than when they were required to interpret the story first ($M = 3.67$, $SD = 1.07$), simple effects test, $F(1,39) = 5.33$, $p < .05$.

Second, to determine whether the two age groups differed in reading time, a 2 (age group) × 2 (story) analysis of variance (ANOVA) was conducted. Neither age group nor the Age Group × Story interaction was significant. There was, however, a significant main effect for story, $F(1,41) = 5.76$, $p < .05$. Participants who read the Stream story took more time ($M = 150.59$ s, $SD = 55.24$) than did those who read the Monk story ($M = 118.78$ s, $SD = 30.59$).

Third, to determine if any sex differences existed prior to the age comparative analysis, separate Sex × Age Group ANOVAs were conducted on the following measures: the four story-rating scales, the propositional recall scores, and the two interpretation scores. None of the analyses produced significant sex effects or Sex × Age Group interactions. All subsequent age comparisons were thus collapsed across the sex variable.

Age-Group Differences in the Retelling Task

To examine age-group differences in recall for the literal propositional text, the number of propositions produced in the retelling protocols was first transformed into a proportion score. Then, to simplify the gist-detail analysis, the original raw scores for Importance Levels 1 and 2 (main ideas) were summed to form a single gist level score, and Importance Levels 3 and 4 (details) were summed to form a single detail level score. Next, the gist score was transformed into a proportion using the sum of the total number of propositions in Levels 1 and 2 as the denominator. The detail score was likewise transformed using as a denominator the sum of the total number of propositions in Levels 3 and 4. A 2 (age group) × 2 (response task order) × 2 (story) × 2 (importance level) ANOVA with repeated measures on the last factor was conducted on the proportion data.

Total recall. — The main effect for age group was significant, $F(1,39) = 13.79$, $p < .001$, but qualified by a significant interaction between age group and response task order, $F(1,39) = 11.34$, $p < .01$. As shown in Figure 1, the age-group effect was obtained in the retell-first but not in the interpret-first condition, simple effects test, $F(1,39) = 18.93$, $p < .001$. When asked to retell the story first, younger adults recalled more of the propositional text ($M = .37$, $SD = .04$) than older adults ($M = .23$, $SD = .08$). By contrast, when asked to interpret the story first, there were no age differences in propositional recall. Note that in this condition, compared to the retell-first condition, the younger adult mean score dropped about 6% ($M = .31$, $SD = .09$), and the older adult mean score increased about 5% ($M = .28$, $SD = .12$).

There was no main effect for response task order nor were there any other significant interactions. There was, however, a significant main effect for story, $F(1,39) = 6.23$, $p < .05$. In general, propositional recall was higher for the Monk story ($M = .32$, $SD = .08$) than for the Stream story ($M = .27$, $SD = .11$).

Gist versus detail recall. — In the same analysis, the main effect for importance level was significant, $F(1,39) = 430.81$, $p < .001$, while the critical Age Group × Importance Level interaction was nonsignificant. Thus, both younger and older adults were sensitive to a text's hierarchical structure, recalling more of a story's gist than its details. Although the Story × Importance Level and the Age Group × Story × Importance Level interactions were significant, $Fs(1,39) = 9.89$ and 9.37, respectively, $ps < .01$, this was simply due to the fact that there were significant age differences at both levels in the Stream story but not in the Monk story. For the Monk story, older adults recalled as much of the gist as did the younger adults.

Age-Group Differences in the Interpretation Task

Length of interpretation. — To examine age-group differences in the length of the interpretation responses, a 2 (age group) × 2 (response task order) × 2 (story) ANOVA was conducted on the number of words in the responses. We obtained significant main effects for age group, $F(1,39) = 6.29$, $p < .05$, and response task order, $F(1,39) = 7.03$, $p < .05$, but not for story. In general, the interpretations of the younger adults were longer ($M = 245.00$ words, $SD = 176.35$) than those of the older adults ($M = 148.75$, $SD = 84.62$), and longer interpretations were produced in the retell-first ($M = 245.96$, $SD = 180.19$) compared with the interpret-first condition ($M = 143.57$, $SD = 63.16$). A marginally significant Age Group × Response Task Order interaction, $F(1,39) = 3.83$, $p = .058$, qualifies these main effects: The interpretations of the younger adults who retold the story first were over twice as long ($M = 329.75$,

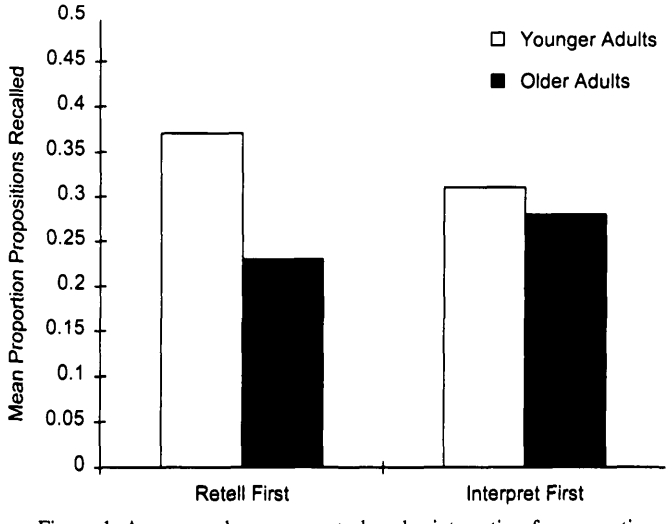

Figure 1. Age group by response task order interaction for proportion of propositional recall.

SD = 201.55) as those of the younger adults who interpreted the story first (M = 152.55, SD = 75.60), simple effects test, $F(1,39) = 10.43$, $p < .01$. There were no significant differences between the length of the interpretations produced by the older adults within either response task condition.

Quality of interpretation. — To arrive at a single score for the depth and synthesis scales, the mean score of the coders for each scale was calculated. The resultant scores ranged from 1 to 5 for both scales, with a mean of 3.04 (SD = 1.00) for the depth scale and a mean of 2.94 (SD = 1.12) for the synthesis scale. Depth and synthesis were correlated among older adults (r = .67, $p < .001$) and younger adults (r = .64, $p < .01$). Depth was not associated with interpretation length in either age group. Synthesis scores were negatively related to interpretation length among the younger adults (r = –.64, $p < .01$) but not among the older adults.

Because we were interested in age differences in the qualitative nature of the interpretations, we needed to identify individual patterns in the interpretive response protocols. The process of identifying patterns began by recoding the scale scores. Specifically, each scale was dichotomized such that values at or above 2.50 (see Appendix) represented a high rating and values below 2.50 represented a low rating. Accordingly, ratings on the depth scale at or above 2.50 were designated as deep responses, and those below 2.50 were designated as shallow responses. Similarly, ratings on the synthesis scale at or above 2.50 were designated as high synthetic responses, and those below were designated as low synthetic (or analytic) responses.

This procedure yielded four patterns of interpretive responding. These patterns are shown in Table 2 with their corresponding means on the two rating scales. Pattern 1, with high values on both the depth and synthesis scales, was designated as the *deep-synthetic* style. These interpretations represented the deeper meanings of a story in a succinct and highly integrative way. Pattern 2, with low values on depth and high values on synthesis, was called *shallow-synthetic*. These interpretations were global, surface-level kinds of interpretations, much like a summary or the gist of a story. Patterns 3 and 4 were both analytic, with low values on the synthesis scale. Pattern 3, with high values on depth, reflected the deeper symbolic themes of the story and was called *deep-analytic*. By contrast, Pattern 4 interpretations, with low depth values, were more literal and were called *shallow-analytic*.

To examine the relationship between age and interpretation pattern, a chi-square analysis was conducted. The number of participants producing each interpretive pattern by age group is shown in Table 3. The chi-square was significant, χ^2 (3, N = 47) = 8.46, $p < .05$. As shown in the table, more older than younger adults produced deep-synthetic interpretations of a story, and more younger than older adults produced analytic interpretations, either deep or shallow. In addition, within the older adult group, there was a markedly greater incidence of deep-synthetic interpretations than of analytic interpretations. By contrast, the younger adults' interpretations were more evenly split between the deep-synthetic and analytic patterns. However, like the older group, the younger adults favored deep (n = 14) over shallow (n = 9) responding. Finally, it is important to note that the shallow-synthetic (or global gist) pattern was produced by only five individuals and was unrelated to age.

We also examined the association between interpretation pattern and response task order. The chi-square was nonsignificant. Although not tested due to the small cell sizes, the Age Group × Interpretation Pattern × Response Task Order associations are interesting to note. Specifically, 8 of the 12 younger adults who produced analytic interpretations were in the retell-first condition. Among the nine younger adults with deep-synthetic patterns, seven interpreted the story first and only two retold the story first. By contrast, the influence of response task order was not as pronounced in the older adult group. Ten of the 17 older adults with deep-synthetic responses were in the interpret-first condition, whereas 7 were in the retell-first condition. Finally, there was no significant association between the story read and the pattern of interpretation produced.

To further distinguish the four interpretation patterns on the two criteria, we ran separate one-way ANOVAs for depth and synthesis. Following each ANOVA, post hoc Duncan Multiple Range Tests (set at .05) were conducted. For the depth criterion, the one-way ANOVA was significant, $F(3,46) = 45.31$, $p < .001$, suggesting mean differences in depth among the four patterns (see Table 2). The post hoc analysis showed that the Pattern 1 mean was significantly greater than the means on the other three patterns; the Pattern 3 mean was greater than the means for Patterns 2 and 4, which were not significantly different from each other. For the synthesis criterion, the one-way ANOVA was also significant, $F(3,46) = 29.96$, $p < .001$. Again, the Pattern 1 mean distinguished itself from the other three pattern means; the Pattern 2 mean was greater than the means for Patterns 3 and 4, which were not significantly different from each other. Thus, Pattern 1, the preferred style of the older adults, was clearly distinguished from the others as the deepest and most synthetic.

Finally, to see whether the age groups differed, on average, on either of the two criteria, separate 2 (age group) ×

Table 2. Number of Interpretive Responses Identified by Pattern and Their Corresponding Means on the Two Rating Criteria

		Depth		Synthesis	
Pattern	n	M	SD	M	SD
Pattern 1: Deep-synthetic	26	3.76	0.60	3.71	0.78
Pattern 2: Shallow-synthetic	5	2.20	0.18	2.87	0.43
Pattern 3: Deep-analytic	5	2.94	0.44	1.85	0.32
Pattern 4: Shallow-analytic	11	1.76	0.37	1.67	0.48

Table 3. Number of Participants Producing Each Interpretive Pattern by Age Group

Age Group	Deep-synthetic	Shallow-synthetic	Deep-analytic	Shallow-analytic
Younger adults	9	2	5	7
Older adults	17	3	0	4

2 (response task order) × 2 (story) ANOVAs were run for the depth and the synthesis scale scores. As suggested by Keppel (1991), Bonferonni tests were utilized by setting the critical alpha at .025. There were no significant age-group differences on the depth criterion (younger group $M = 2.88$, $SD = 1.01$; older group $M = 3.18$, $SD = 0.99$). However, the age-group difference for the synthesis criterion approached significance, $F(1,39) = 5.20$, $p = .028$, with the older group's interpretations rated higher or, on average, more synthetic ($M = 3.30$, $SD = 1.08$) than the younger group's ($M = 2.57$, $SD = 1.06$).

Discussion

The results of this study support the expected pattern of age-related differences in performance on the retell and interpretation tasks. When asked to retell a story as close to the original as possible, the younger adults recalled more propositions than did the older adults in the retell-first, although not in the interpret-first, condition. In addition, both older and younger adults recalled more of a story's main ideas (gist) relative to its details. When asked to interpret the same story, most members of both age groups focused on the story's deeper meanings. More older than younger adults, however, produced deep and synthetic interpretations. In addition, there was a clear preference among the older age group for deep-synthetic responding. Although more younger than older adults produced analytic (deep or literal) interpretations, within the younger group there was no clear preference for either an analytic or a deep-synthetic style.

In and of themselves, our retell task results are not surprising, as they are consistent with most studies of age differences in text recall (for review, see Zelinski & Gilewski, 1988). But when examined within the more inclusive context of the interpretation task, an expanded picture of older adults' text-processing capabilities emerges. Our results suggest that it is not simply the gist of a story that is encoded in older adults' memories but also the story's deep interpretive meanings. Moreover, the interpretive meanings appear to be encoded into patterned wholes. It is important to keep in mind that the deep-synthetic interpretations observed in this study were not merely reflections of cursory and global encodings, as we were able to distinguish such gistlike responses (our shallow-synthetic pattern) from the deeper and more integrated responses. Indeed, the deep-synthetic pattern was distinguished from the other patterns, as it was rated significantly highest on both depth and synthesis criteria.

The deep-synthetic interpretations produced by most of the older adults are not viewed as "better" than the analytic interpretations produced by many of the younger adults. As stated previously, both age groups showed an ability to focus on the deeper-level meanings in a story. Rather, we view the difference to be in the use of qualitatively different strategies. Although both literal and interpretive aspects of a story are probably simultaneously encoded by all readers, younger adults seem more likely than older adults to focus on encoding the details of the text at both literal and interpretive levels. Older adults seem to take a different approach, focusing less on encoding details and more on encoding the gist of a story as well as its underlying significance.

That older adults focus on the depth and synthesis of narrative meaning is consistent with the idea that abstract thematic knowledge structures — representing the point, significance, or even moral of a story — are built up over one's lifetime as numerous individual episodes involving related experiences are connected, organized, and stored in memory (Black & Seifert, 1985; Dyer, 1983). Because older adults have accrued a lifetime of experiences, such structures in memory may serve to guide their processing toward a search for the underlying themes in narrative.

Such a mode of processing may well have an adaptive value for the aging cognitive system. There are two possible ways to account for such an adaptive change. On one hand, it may be how the aging mind compensates for some loss in the ability to efficiently process more detailed information (see Baltes, 1987, 1993). On the other hand, it may be the result of a cognitive progression in adulthood toward encoding and storing new information into highly integrative units of meaning — units that carry far greater temporal stability than the fragile propositional units on which they are based (see Labouvie-Vief & Schell, 1982). By its very nature, then, a highly integrative or synthetic approach is likely to involve the "loss" of details. Loss in this sense, however, is interpreted in the context of cognitive "growth" (see Labouvie-Vief, 1990) rather than in the context of decline with compensation.

The fact that the age effect in the retell task was significant in the retell-first but not in the interpret-first condition requires some discussion. When asked to interpret the story before retelling it, the younger adults lost their recall advantage over the older adults. This loss in advantage was due, in part, to a decrease in the propositional recall of the younger adults and, in part, to an increase in the older adults' propositional recall. Why might this be so?

One possible explanation is that interpreting a story first may have served to disrupt the subsequent recall of the younger but not older adults, causing the younger adults to lose some detail that they were later unable to retrieve. For the older adults, the situation may have been different. Recall that the older adults who interpreted a story first rated it to be less difficult and more familiar than did their age counterparts who retold a story first. Perhaps the initial production of a deep and synthetic interpretation served as an elaborate memory structure for this age group, facilitating the retrieval of story detail.

An alternative explanation is that the younger adults were simply more biased by initial response task demands than were the older adults. If the initial response task was to retell a story, then the younger adults may have focused on the propositions and a lengthy reproduction of the story. If, however, the initial response task was to interpret the story, their focus may have been less on the details and more on abstract themes.

This explanation is consistent with the interpretation data. When asked to interpret a story before retelling it, more younger adults produced deep-synthetic responses than analytic responses; when asked to retell the story first, their interpretations were more often analytic and were lengthier as well. As college students, it is likely they have learned the value of being responsive to the changing

demands of different test-taking situations. By contrast, most older adults produced deep-synthetic interpretations regardless of task order. That their interpretive responding did not seem to be influenced by task order lends further support to the idea that deep-synthetic processing is the more natural strategy associated with age.

This study raises a number of issues needing further examination. First, although our results generalized across stories, lending credence to their robustness, both stories were Sufi teaching tales. It remains to be seen whether the same effects are observed with other kinds of narrative texts. Second, because our older sample was drawn from a highly educated population, the age-related trend toward synthetic processing observed in this study may only hold true for well-educated older adults. Whether the same trend would hold true for less educated populations remains to be examined. Third, the observed age-related patterns raise important, yet unanswered, questions regarding the match (or mismatch) between a particular cognitive task and the social-cognitive roles and skills associated with an individual's life stage. The retell and interpretation tasks may have differentially tapped the social-cognitive goals associated with two very different life stages — the retell task being a better match for the goals of youth and the interpretation task better matched to the goals of later life. As several authors have pointed out, youth is the period in life when it is important to efficiently acquire, retain, and recall large bodies of new information; the young cognitive system seems well prepared to do so (Labouvie-Vief, 1985; Mergler & Goldstein, 1983; Schaie, 1977/1978). By contrast, in later life it may become more important to effectively access and apply the extensive knowledge and wisdom one has acquired through a lifetime of experiences (Baltes, 1993; Labouvie-Vief, 1985; Mergler & Goldstein, 1983; Meyer, 1986; Perlmutter, 1988). Thus, tasks that draw on the application of experiential knowledge and on the ability to extract what is essential (such as the interpretation task) may be more in line with the processing goals and skills of later life. In future research on cognition and memory in adulthood, the degree to which the demands of the task match the social-cognitive goals of the learner need to be examined (see also Blanchard-Fields & Abeles, 1996).

Finally, we are left with the question of whether older, like younger, adults can reproduce (from memory) lengthy and detailed accountings of complex narrative texts. And, if so, under what circumstances are they likely to do so? To put this question to the test, we believe it is necessary to manipulate the learning and recall contexts in ways that — unlike standard laboratory conditions — *naturally* demand the construction and accurate representation in memory of a more complete propositional text. Consequently, we are currently examining the influence of young children (compared with adult researchers) as the listeners in story retelling tasks.

In summary, the results of this study extend our understanding of the text-processing capabilities of older adults and contribute to the growing literature on qualitative age differences in narrative recall. We have shown that later life processing is probably more than a simple trade-off between details and gist. Whether one adopts a "loss with compensation" or a "loss due to growth" interpretation of our data, it is becoming clear that age and experience bring the increasing ability not only to focus on what is important but also to tie together the deep elements of narrative succinctly and completely.

ACKNOWLEDGMENTS

This research was supported by a NIA training grant provided to the University of Michigan Institute of Gerontology. We wish to thank two anonymous reviewers for their helpful comments; Adaline Adams, Lisa Conn, Ada Cowan, David Craig, Peggy Gardner, Keith Markman, Bruce Moore, Lisa Ribiat, Tom Spitzer, Tom Sullivan, Norbert Vnek, and Myra White for their assistance in the propositional and interpretation coding; and Loretta Vitolo for her editorial assistance.

The Editor thanks Dr. Anderson D. Smith, Georgia Institute of Technology, who served as Review Editor for the manuscript.

Address correspondence to Dr. Cynthia Adams, Center for the Study of Women in Society, 340 Hendricks Hall, 1201 University of Oregon, Eugene, OR 97403-1201.

REFERENCES

Adams, C. (1991). Qualitative age differences in memory for text: A life-span developmental perspective. *Psychology and Aging, 6,* 323–336.

Adams, C., Labouvie-Vief, G., Hobart, C. J., & Dorosz, M. (1990). Adult age group differences in story recall style. *Journal of Gerontology: Psychological Sciences, 45,* P17–P27.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology. *Developmental Psychology, 23,* 611–626.

Baltes, P. B. (1993). The aging mind: Potential and limits. *The Gerontologist, 33,* 580–594.

Birren, J. E. (1969). Age and decision strategies. In A. T. Welford & J. E. Birren (Eds.), *Interdisciplinary topics in gerontology* (Vol. 4, pp. 23–36). New York: Karger.

Black, J. B., & Seifert, C. M. (1985). The psychological study of story understanding. In C. R. Cooper (Ed.), *Researching response to literature and the teaching of literature: Points of departure* (pp. 190–211). Norwood, NJ: Ablex.

Blanchard-Fields, F., & Abeles, R. P. (1996). Social cognition and aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 150–161). San Diego, CA: Academic Press.

Boswell, D. A. (1979). Metaphoric processing in the mature years. *Human Development, 22,* 373–384.

Brodman, K., Erdman, A. J., & Wolff, H. G. (1949). *Cornell Medical Index Health Questionnaire.* Ithaca, NY: Cornell Medical College.

Bruner, J. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Carstensen, L., & Turk-Charles, S. (1994). The salience of emotion across the adult life span. *Psychology and Aging, 9,* 259–264.

Cohen, G. (1988). Age differences in memory for text. Production deficiency or processing limitations? In L. L. Light & D. M. Burke (Eds.), *Language, memory, and aging* (pp. 171–190). New York: Cambridge University Press.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer.* New York: Free Press.

Dyer, M. G. (1983). *In-depth understanding: A computer model of integrated processing for narrative comprehension.* Cambridge, MA: MIT Press.

Fry, E. (1968). A readability formula that saves time. *Journal of Reading, 11,* 513–516, 575–578.

Hartley, J. T. (1989). Memory for prose: Perspective on the reader. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 135–156). New York: Cambridge University Press.

Hollingshead, A. B. (1975). *Four factor index of social status.* Unpublished manuscript. New Haven, CT: Yale University.

Holyoak, K. J. (1982). An analogical framework for literary interpretation. *Poetics, 11,* 105–126.

Hultsch, D. F., & Dixon, R. A. (1984). Memory for text materials in adulthood. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 6, pp. 77–108). New York: Academic Press.

Hultsch, D. F., & Dixon, R. A. (1990). Learning and memory in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 258–274). San Diego, CA: Academic Press.

Jastak, J. F., & Jastak, S. R. (1964). Short forms of the WAIS and WISC vocabulary subtests. *Journal of Clinical Psychology, 20,* 167–199.

Jepson, K. L., & Labouvie-Vief, G. (1992). Symbolic processing of youth and elders. In R. L. West & J. D. Sinnott (Eds.), *Everyday memory and aging* (pp. 124–137). New York: Springer-Verlag.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

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

Kogan, N., Connor, K., Gross, A., & Fava, D. (1980). Understanding visual metaphor: Developmental and individual differences. *Monographs of the Society for Research in Child Development, 45*(1, Serial No. 183).

Kramer, D. A., & Kahlbaugh, P. E. (1989). *Memory for a dialectical and a nondialectical prose passage in young and older adults.* Unpublished manuscript. New Brunswick, NJ: Rutgers–The State University of New Jersey.

Labouvie-Vief, G. (1985). Intelligence and cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 500–530). New York: Van Nostrand Reinhold.

Labouvie-Vief, G. (1990). Modes of knowledge and the organization of development. In M. L. Commons, C. Armon, L. Kohlberg, F. A. Richards, T. A. Grotzer, & J. Sinnott (Eds.), *Adult development: Models and methods in the study of adolescent and adult thought* (Vol. 2, pp. 43–62). New York: Praeger.

Labouvie-Vief, G., & Schell, D. (1982). Learning and memory in later life. In B. Wolman & G. Stricker (Eds.), *Handbook of developmental psychology* (pp. 828–846). Englewood Cliffs, NJ: Prentice-Hall.

Mergler, N. L., & Goldstein, M. D. (1983). Why are there old people? Senescence as biological and cultural preparedness for the transmission of information. *Human Development, 26,* 72–90.

Meyer, B. J. F. (1986). Reading comprehension and aging. In K. W. Schaie & C. Eisdorfer (Eds.), *Annual review of gerontology and geriatrics* (pp. 93–115). New York: Springer.

Meyer, B. J. F., & Rice, G. E. (1989). Prose processing in adulthood: The text, the reader, and the task. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 157–194). New York: Cambridge University Press.

Perlmutter, M. (1988). Cognitive potential throughout life. In J. E. Birren & V. L. Bengtson (Eds.), *Emergent theories of aging* (pp. 247–268). New York: Springer.

Perlmutter, M., & Nyquist, L. (1990). Relationships between self-reported physical and mental health and intelligence performance across adulthood. *Journal of Gerontology: Psychological Sciences, 45,* P145–P155.

Schaie, K. W. (1977/1978). Toward a theory of adult cognitive development. *International Journal of Aging and Human Development, 8,* 129–138.

Shah, I. (1967). *Tales of the dervishes.* London: Jonathan Cape.

Stine, E. A. L., & Wingfield, A. (1990). The assessment of qualitative age differences in discourse processing. In T. M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 35–92). Amsterdam: North-Holland.

Turner, A., & Greene, E. (1977). *The construction and use of a propositional text-base* (Tech. Rep. No. 63). Boulder: University of Colorado, Institute for the Study of Intellectual Behavior.

White, E. M. (1985). *Teaching and assessing writing.* San Francisco: Jossey-Bass.

Zelinski, E. M., & Gilewski, M. J. (1988). Memory for prose and aging: A meta-analysis. In M. L. Howe & C. J. Brainerd (Eds.), *Cognitive development in adulthood: Progress in cognitive development research* (pp. 133–158). New York: Springer-Verlag.

Received March 11, 1996
Accepted January 27, 1997

Appendix

To identify individual patterns in interpretive response protocols, each scale was dichotomized so that values at or above 2.50 represented a high rating and values below 2.50 represented a low rating. This dichotomization was applied at approximately the scale midpoint to identify interpretations that were truly high and low on each criterion. Note that although the scale midpoint was 3, each rating represents the mean of three or four coders. Thus, using a cutoff value of 2.50 indicates that the majority of coders rated the interpretation as high (or low) on the criterion, and the properties of the coding decisions determined the split. In contrast, the distribution of scores across the criterion scale would determine high and low groupings if a median split was used. Careful training of the coders and the integrity of the coding led us to believe that those interpretations scored at or above (or below) the scale midpoint were indeed high (or low) on that scale.