ABSTRACT—In a multistage experiment, twelve 4- and 9-year-old children participated in a triad rating task. Their ratings were mapped with multidimensional scaling, from which euclidean distances were computed to operationalize semantic distance between items in target pairs. These children and age-mates then participated in an experiment that employed these target pairs in a story, which was followed by a misinformation manipulation. Analyses linked individual and developmental differences in suggestibility to children’s representations of the target items. Semantic proximity was a strong predictor of differences in suggestibility: The closer a suggested distractor was to the original item’s representation, the greater was the distractor’s suggestive influence. The triad participants’ semantic proximity subsequently served as the basis for correctly predicting memory performance in the larger group. Semantic proximity enabled a priori counterintuitive predictions of reverse age-related trends to be confirmed whenever the distance between representations of items in a target pair was greater for younger than for older children.

By most measures, research on children’s suggestibility has been a success story (see recent issues of Applied Cognitive Psychology, Volume 18, Issue 8, 2004, and Developmental Review, Volume 24, Issue 1, 2004). Despite this success, extant developmental theory has made only one unambiguous prediction about children’s susceptibility to suggestion—a main effect for chronological age: Children become less vulnerable to suggestions as they get older (Bruck & Ceci, 1999; Ceci & Bruck, 1993). Although chronological age is almost always the strongest predictor of suggestibility, there is much variability within age groups. Despite a great deal of work in this area, examination of individual differences has not been notably successful in explaining within- or between-age variance in suggestibility (Bruck & Melnyk, 2004; Chae & Ceci, 2005).

In the quest to understand individual differences in suggestibility, we wanted to explore whether underlying knowledge representations shape and constrain processing activity. Specifically, our goal in the current study was to investigate the role of underlying semantic knowledge structures in explaining both individual and developmental differences in suggestibility.

Semantic knowledge has long played a prominent role in many theories of adult memory. For example, spreading-activation theory (Collins & Loftus, 1975), cue-compound theory (Ratcliff & McKoon, 1994), fuzzy-trace theory (Brainerd & Reyna, 2002, 2005a), and the levels-of-processing approach (Craik & Tulving, 1975) all have at their core the finding that semantic processes have an important influence on reaction time, priming, recall, recognition, and false memory. The Deese-Roediger-McDermott (DRM) paradigm has spawned its own cottage industry of research demonstrating semantic influence on false memory (e.g., Roediger & McDermott, 1995). None of these theories, however, were crafted to explain individual or developmental changes in semantic processing and the role of such differences in suggestibility.

In the present study, we developed a conceptual and methodological approach to fill this void. We explored the manner in which children’s mental representations individually change with development and influence susceptibility to suggestion. We first analyzed individual children’s similarity judgments of stimuli in order to model differences in their representations of those stimuli (dimensionality, semantic distance). These distances were next used to predict the children’s vulnerability to suggestion. We then made targeted predictions about other children’s suggestibility for these same stimuli, basing these predictions on each stimulus’s distance from its paired distractor.

Our approach employed multidimensional scaling (MDS) procedures to reveal children’s underlying knowledge structures and to link this representational information to individual and developmental differences in suggestibility. Given previous findings, we expected age to have a main effect on suggestibility.
In addition, we predicted a strong influence of semantic distance on memory performance; specifically, we expected that items closer in representational space would be more vulnerable to suggestion. This prediction was based on an assumption that greater proximity of stimuli in representational space would make a distractor more difficult to distinguish from its paired target. We also predicted a priori reverse age differences in suggestibility in an independent sample of age-mates whenever items were closer in representational space for older children than for younger children. Our findings provide compelling power for explaining individual and developmental differences in suggestibility in terms of underlying representational constraints.

METHOD

Subjects
For Phase 1, six 4-year-olds and six 9-year-olds participated in a triads similarity-rating task. For Phase 2, 62 age-mates (M = 4 years 9 months, SD = 2.2 months; M = 9 years 9 months, SD = 6.2 months) joined these 12 children.

Phase 1 Procedure
For Phase 1, 1,540 triads were systematically generated from the three-way permutations of 22 pictorial stimuli depicting common objects. Each child was individually presented with one sixth of these triads of the full set. Hence, each child rated 257 of the 1,540 triads:

\[
\frac{22!}{(22-3)!(3!)} = 1,540, \frac{1,540}{6} = 257
\]

For each triad rated, a child was presented with three color pictures (each measuring 6.2 cm × 6.7 cm) and asked which picture did not belong with the other two. The child was asked to explain each choice before being presented with the next triad. Children took approximately 10 to 20 s per triad rating. To prevent boredom and exhaustion, we administered the 257 triads to each child during multiple brief sessions over a period of 3 to 5 weeks.

Phase 2 Procedure
One to 3 months after the triad sessions, these same 12 children were joined by 62 age-mates, who were randomly assigned to two conditions, control versus misled. Word lists were counterbalanced, leading to four groups: Control Groups A and B and Misinformation Groups A and B.

Children were individually presented an illustrated story that contained 8 critical objects chosen from the set of 22 stimuli from the Phase 1 triad task. For children assigned to the misinformation groups, 4 of these critical items were later the targets of active false suggestions (e.g., if lemon had been in the story, then orange was suggested). The other four critical items were not the targets of overt false suggestions but were later paired with semantic distractors during a recognition test to examine autosuggestibility.1 Items were depicted visually in the story, and assignment of objects to be critical items or distractors (false suggestions or autosuggestions) was counterbalanced across groups; children in Group A (control and misinformation) were presented with one set of critical items (egg, ant, crab, milk, cow, horse, robin, and lemon), whereas those in Group B were presented with a complementary set of critical items (cheese, spider, lobster, soda, deer, bear, eagle, and orange). Thus, the false suggestions for Group A served as critical items in Group B’s story and vice versa, as a control for specific items’ memorability. Otherwise, the procedures for Groups A and B were identical.

Two days after listening to the story, all children were interviewed about critical and noncritical items presented in the story. Children in the misinformation groups were interviewed with a standard verbal misinformation procedure. In addition to asking about noncritical items, the interviewer falsely suggested a set of four semantically associated distractors in place of four critical items (Ceci, Ross, & Toglia, 1987). For example, in one of the stories, a boy and girl saw an eagle in the zoo’s aviary, but the children were told during the interview that the bird was a robin. Children in the control groups received a similar interviewer, with the exception that it contained no misinformation.

Five to 7 days later, all children were shown 32 pictorial stimuli serially and asked which ones they remembered seeing in the story.2 In addition to the 8 critical items from the story and their 8 distractors (4 false suggestions and 4 semantically related autosuggestions), the test stimuli included 16 unrelated pictures from the story that were not included in the triad-sorting task. Pictures were presented in randomized order.

RESULTS

Individual-Level Analyses
Our first goal was to determine whether children’s semantic representations constrain their suggestibility at the individual level. That is, does semantic distance predict individual children’s memory performance, as evidenced by increased suggestibility for semantically closer items? To answer this question, we analyzed the data from the 12 children who completed the Phase 1 triad task, for whom we could calculate both

---

1 Binet (1900) proposed a form of suggestion he termed autosuggestion. It refers to an endogenous suggestion that emanates internally, not from experimenter-provided misinformation. Brainerd and Reyna (1995) showed that autosuggestibility is a powerful source of false memories in children. In the present experiment, autosuggestions were semantically related distractors that, by themselves, might be suggestive because of the child’s internal reasoning processes.

2 As Brainerd and Reyna (2003b) noted, the standard forced-choice procedure results in a conflation of two opponent processes that cannot be disentangled, namely, an increase in false alarms and a decrease in hits. Our procedure avoids this conflation.
A limitation of MDS is that solutions are influenced by the rated similarities among all of the stimuli being judged. Changes to the total constellation of items being judged can result in fluctuations in semantic distance between items in a given pair because the distance between any two items is derived, in part, from the distances among all of the other items. To the extent that such relativity can result in fluctuations in semantic distance between items in a group of 12 children, one full set by younger children and one by older children. This design allowed for the subsequent computation of group-level semantic distances between items without requiring each child to judge every one of the permutations. For all 12 children and for all models, evaluation of s-stress values and scree plots indicated two-dimensional solutions.

To operationalize semantic proximity between paired items (i.e., original critical item and the false suggestion or autosuggestion), we calculated the euclidean distance between each item and its distractor. These distances were entered as a continuous predictor for each criterion (explained later in this section in the discussion of the logistic regression). Qualitatively, older children appeared to possess more refined dimensions than younger children. In their triad ratings, older children never grouped eagle with robin if the predator bear was the third member of the triad; their verbal justifications reflected this reasoning (saying that bears and eagles are both predators). In contrast, younger children always grouped eagle with robin, regardless of the third member of the triad (giving as their reason that eagles and robins are both birds).3

Recognition data were scored dichotomously according to whether or not the child stated the item was present in the story. Of interest was false recognition of a suggested distractor concomitant with rejection of the original item. This pattern (i.e., the classic misinformation effect) reflects the strongest manifestation of suggestibility and also avoids “yes” response biases because the dependent measure is a combined yes-and-no judgment.

To account for nonindependence in the binary response data, we ran a repeated measures logistic regression using a generalized estimating equation approach (Liang & Zeger, 1986) to predict recognition accuracy from age, euclidean distance between paired items, and their interaction. To address the downward bias in the sandwich variance resulting from our sample size (e.g., Mancl & DeRouen, 2001; Pan & Wall, 2002), we ran all tests after multiplying the covariance matrix by a scalar equal to $J/(J - 1)$, where $J$ is the number of clusters in the estimation sample (see Hardin & Hible, 2003). No differences were found in tests distinguishing distractors as active suggestions versus autosuggestions; therefore, all models were collapsed across this variable.

Results were consistent with the suggestibility literature: As age increased, recognition of the distractor in the absence of recognition of the original item decreased (i.e., a suggestibility effect). $\chi^2(1, N = 12) = 7.56, p = .006, p_{rep} = .97$. As predicted, as distance between critical and suggested items increased, false memory of the suggested distractor decreased. $\chi^2(1, N = 12) = 8.53, p = .004, p_{rep} = .98$. Thus, a semantically proximal distractor had a far more deleterious effect than a more semantically distal item. The effects for distance and age did not interact.

Given that, after controlling for age, semantic distance was a highly significant predictor of suggestibility, we next tested the relative importance of these two measures. We were interested in this comparison because the literature indicates that chronological age is the best predictor of suggestibility. Because standardized estimates are not available in the repeated measures logistic regression approach, we plotted receiver-operating characteristic (ROC) curves for the separate models predicting suggestibility from age and distance.4 As shown in Figure 2, there was no significant difference between the areas under the two curves, $\chi^2(1) < 1$; the areas under the ROC curves for age and distance were .640 and .678, respectively, both statistically significant, $p = .023 (p_{rep} = .94)$ and $p = .004 (p_{rep} = .98)$. There was no significant difference when the areas under the two curves were compared with the area under the ROC for the full model (area = .674, $p = .005$), $\chi^2(2) = 1.34, p = .51$. Thus, semantic distance appears to be at least as potent a factor as age in predicting suggestibility.

---

3A limitation of MDS is that solutions are influenced by the rated similarities among all of the stimuli being judged. Changes to the total constellation of items being judged can result in fluctuations in semantic distance between items in a given pair because the distance between any two items is derived, in part, from the distances among all of the other items. To the extent that such relativity operated in the present study, it would have made our findings more robust.

4ROC curves represent the accuracy of a test as the area under the curve derived by plotting the proportion of true positives (sensitivity) against the proportion of false negatives (1 − specificity; Hanley & McNeil, 1982). Relative performance of two tests can be evaluated by comparing areas under the two curves.
Group-Level Analyses

Although the differences in the predictive values of age and semantic distance were statistically nonsignificant, an intriguing finding from the individual-level analyses concerned the directionality of the results; it appeared that semantic distance superseded age as a predictor of suggestibility. Therefore, we tested if we could predict the patterns of suggestibility in our full sample from the semantic-proximity findings from the individual-level MDS solutions. That is, given that distance was at least as good as age at predicting the misinformation effect at the individual level, could group-level differences in one sample be used to predict suggestibility in a larger sample? It would be particularly interesting if some distances between original critical items and their distractors were smaller for the older children than for the younger children, leading to the strongest type of developmental prediction—reverse age-related trends, with older children being more suggestible than younger children for these items. These item-level predictions would be impossible using age alone to predict memory performance.

To determine representative distance at each age for each pair of items, we ran separate weighted MDS (INDSCAL) models (Carrol & Chang, 1970) for younger and older children. These analyses produced common MDS solutions, as seen in Figure 3. Coordinates from the INDSCAL solutions were used to calculate group-level euclidean distances between paired items (see Table 1). We reran each of these INDSCAL analyses, removing children one at a time on the basis of their weirdness factor, an indication of their disagreement with the solutions of their same-age counterparts. All of these reduced analyses showed only slight differences from the analyses for the full sample.5

Fig. 2. Receiver-operating characteristic curves for models predicting children’s suggestibility from (a) age and (b) euclidean distance.

Fig. 3. Nonmetric multidimensional scaling solutions derived from INDSCAL models for the (a) younger and (b) older children’s similarity ratings.

5We also ran a series of Mantel tests to examine the stability and reliability of the results. The INDSCAL distance matrix for younger children’s group data correlated much more positively with the individual matrices from the younger children’s data than with the individual matrices from the older children’s data (four significant correlations out of six vs. three nonsignificant correlations and two significant negative correlations with older children’s individual matrices). Likewise, the INDSCAL distance matrix for older children’s group data correlated more highly with older than with younger children’s individual matrices (although only slightly). For only 14% of the pairs were the distance matrices of older and younger children significantly correlated. These results suggest that the findings are stable and internally reliable.
Table 1 also presents the group differences in suggestibility predicted on the basis of the euclidean distances. Our a priori prediction was that the group with the smaller distance between the critical and distractor items would be more likely to falsely recognize the distractor in the absence of recognizing the critical item. To make our predictions, we ranked the distances between critical and distractor items for each age group. Pairs that were within one rank of each other were regarded as equivalent for the two age groups; otherwise, we predicted greater suggestibility for the age group that showed the smaller distance for a given pair.

As shown in Table 1, we predicted that younger children would show greater suggestibility for four pairs. For one pair, distances were equivalent, and thus we predicted no age difference in suggestibility. Most notably, however, three pairs were closer in semantic space for older than for younger children, and for these pairs we predicted reverse effects of age, with older children being more suggestible than younger children.

To test these predictions, we grouped items on the basis of the predicted direction of the age effect. Thus, one group included the items for which we predicted younger children would show greater suggestibility, and the other group included the items for which we predicted older children would show greater suggestibility (the one pair for which we predicted no age difference was not included in this analysis). For each item group, we calculated the proportion of suggested items that each child recalled.

A 2 (age) × 2 (item group) × 2 (condition) repeated measures analysis of variance with item group as a within-subjects factor and age and condition as between-subjects factors was conducted on the arcsine-root-transformed proportions. This analysis revealed the expected Item Group × Age interaction, $F(1, 70) = 31.10, p < .001, \eta^2_p = .39, \eta^2 = .31$. Subsequent analyses revealed a highly significant effect of age for items predicted to show more suggestibility among younger children. For these items, false recognition was greater among younger children ($M = .28, SD = .20$)—values back-transformed for interpretability—than among older children ($M = .33, SD = .09$).

A significant effect of age, $F(1, 70) = 3.85, p = .05, \eta^2_p = .05$, was also found for items predicted to show more suggestibility among older children. In this case, older children falsely recognized a greater proportion of the items ($M = .59, SD = .29$) than did younger children ($M = .29, SD = .20$) to test for the possibility that data from the children who completed the MDS triad task were driving these results, we reran these analyses excluding the 12 children from Phase 1. These analyses yielded identical results.

Table 1 also presents the percentage of children in each age group who falsely recalled each distractor. These results are striking. Seven of the eight predictions of group differences based on the euclidean distances were confirmed, including two reverse effects of age. Only the cow-deer pairing did not follow the predicted pattern. Assuming a null probability of .5 for each prediction, the one-tailed $p$ for correctly predicting seven out of eight outcomes equals .035.

**DISCUSSION**

The present study was designed to test a set of interrelated hypotheses about the relation between knowledge representation and suggestibility. We reasoned that a potentially potent individual difference in memory and suggestibility is the way children represent their semantic knowledge. We hypothesized that when distractors are represented near critical items in semantic space, they have the most damaging effect by making the distractors hard to distinguish from the critical items. Although other researchers have argued that semantic knowledge plays a vital role in memory processing and in the development of false memories (e.g., Collins & Loftus, 1975; Craik & Tulving, 1975; Roediger & McDermott, 1995), to our knowledge, no study has...
yet examined the role of semantic knowledge in individual differences in suggestibility, and the sole study to date to have suggested that developmental differences in knowledge may drive age differences in suggestibility was that of Elischberger (2005). He showed that when levels of social and academic knowledge were crossed with the consistency between postevent suggestions and knowledge, suggestions that were consistent were incorporated into false reports more often than suggestions that were contradictory with knowledge. In addition, Park, Shobe, and Kihlstrom (2005) have shown that “coordinate” items, which are the type we employed (i.e., items at the same categorical level, such as lemon, orange, and grapefruit), are 3 times more likely to produce false memories in adults than are items that are in a hierarchical or taxonomic relationship (e.g., fruit, apple, orange).

Our results provide powerful support for the hypothesis that semantic knowledge constrains suggestibility. Euclidean distance not only was a strong predictor that did not interact with age, but seemed to rival age in its predictive potency. Thus, a given child was more likely to confuse a critical item with its distractor the more closely in semantic space that child represented the distractor and the critical item. In fact, euclidean distance was such a strong predictor of suggestibility that we were able to make a priori predictions of reverse effects of age on suggestibility within an independent sample of children. We predicted that older children would be more suggestive than younger children when a critical item and its distractor were represented more closely in semantic space by older children than by younger children, and that the reverse would be true when the critical item and its distractor were represented more closely by younger children. This counterintuitive prediction was supported in seven out of eight cases, and these results lend credence to the view that individual differences in underlying representations play a fundamental role in suggestibility (Ornstein & Elischberger, 2004).

These results also suggest that differences in semantic representations may underlie developmental differences in suggestibility. Classic models of development posit that early development proceeds from syncretic, undifferentiated structures to more distinguishable ones (Werner, 1948; Werner & Kaplan, 1952). The present study provides some evidence that this pattern of increasing differentiation may characterize the developmental changes in children’s representations. As noted, older children’s representations appeared to possess a predacry dimension that younger children’s representations lacked; similarly, it appeared that older children possessed a citrus dimension (lemon, orange, grapefruit) and a dairy dimension (milk, eggs, cheese, butter), whereas younger children depended more on functional and perceptual dimensions (e.g., things to drink, small things, yellow things). It is important to note that when the underlying representations were comparable for older and younger children, we found a marked reduction in age differences in suggestibility.

Until now, chronological age has been the single largest source of variance in suggestibility studies; the effects of no other variable have approached the effects of age in magnitude (Ceci & Bruck, 1993). However, the mechanisms driving this age effect were either unspecified or only hinted at (e.g., source-monitoring changes) and failed to explain much of the developmental and within-age variance (Siegler, 2004). The present study has identified a factor that explains within-age variance in suggestibility and partially explains developmental trends in suggestibility. Still, it is clear that items’ representations are not the only factor driving developmental differences. Although more work will be necessary to flesh out this theoretical approach, the present results hold promise of revealing a powerful mechanism underlying individual differences and developmental trends in the literature, and for making precise predictions regarding within-age and between-age differences.

REFERENCES


(RECEIVED 3/28/06; REVISION ACCEPTED 11/19/06; FINAL MATERIALS RECEIVED 11/30/06)