Developmental Psychology
2006, Vol. 42, No. 6, 1193–1205

0012-1649/06/$12.00 DOI: 10.1037/0012-1649.42.6.1193

Developmental Invariance in Distinctiveness Effects in Memory

Mark L. Howe
Lancaster University

Improvements in 5- and 7-year-olds' acquisition and retention of related concept pairings were examined when additional similarities and differences between pair members were provided. Using a standard paired-associate learning paradigm, children learned 18 related picture pairs; some of the children either were given or produced additional similarities or differences between pair members at the time of learning. Three weeks after learning was complete, children attempted to recall the pairs. Using a model to determine the storage and retrieval loci of these effects, the results showed that (a) all children benefited from self-generated elaborations, regardless of whether these were similarities or differences, and these benefits were storage related, and (b) difference elaborations improved children's retention regardless of whether they were self- or experimenter-generated, and these effects were primarily retrieval based. These results are consistent with theories that (a) view retrieval as the locus of distinctiveness effects and (b) view storage as the locus of self-generated memory improvements.

Keywords: distinctiveness effects in memory, memory development, long-term retention, mathematical modeling

Interest in children's event memory has been fueled by both theoretical and pragmatic concerns. Pragmatically, researchers have been concerned about the accuracy and durability of children's memory reports in forensic investigations. Theoretically, investigators have been concerned about factors underlying changes in children's memory, particularly the fundamental cognitive and neurological processes that are the basis for memory development. Much of the resulting research has shown that very young children can encode, store, and retrieve information about naturally occurring and experimentally contrived events and, under optimal conditions, can retain that information fairly accurately over a number of months or even years (for a review, see Howe, 2000). Evidence is accumulating that many of the variables that govern storage and retrieval in young children are the same as those that regulate these memory processes in older children and adults (for reviews, see Bauer, 2004; Howe, 2000). This continuity of the human memory system notwithstanding, there do exist significant developmental advances in memory across childhood (e.g., faster encoding, storage, and retrieval; greater ability to maintain information in storage for longer periods of time) that are contingent on advances in related cognitive processes (e.g., knowledge, attention, strategies, metamemory; see Bauer, 2005; Howe, 2000).

In spite of this remarkable progress in memory development research, there exist a number of unanswered questions. Foremost among these is the question of the types of events (e.g., traumatic or nontraumatic) that young children remember best and which of these endure over time to become part of the adult repertoire. Recently, Howe (2006) suggested that memory is enhanced not by the positive or negative valence of the event per se but by its distinctiveness. Although numerous definitional issues exist (for a discussion, see Schmidt, 1991), distinctiveness (at least at an operational level) involves the processing of differences relative to some context (Medin, Goldstone, & Gentner, 1993). Distinctive events (or distinctive aspects of events) are those that stand apart from the background of one's previous experiences and knowledge and thus tend to be well remembered over time. Indeed, retrospective research on college students' memory for distinctive autobiographical events indicates that certain of these were retained from the age of about 2 years (Eacott & Crawley, 1998; Usher & Neisser, 1993). There is also considerable evidence from the adult literature that memory, particularly autobiographical memory, is enhanced by event distinctiveness and that what one recalls about one's life are events that are personally consequential (e.g., Conway, 1996), are turning points or formative periods (e.g., Rubin & Schulkind, 1997), or are distinctive with respect to one's self (e.g., Csikszentmihalyi & Beattie, 1979).

The importance of distinctiveness for memory has been amply supported by the results of laboratory research with adults (e.g., Hunt, 2006). In general, it is well known that unique or unusual information tends to be better retained than commonplace information, regardless of whether researchers are measuring recollection in episodic or autobiographical memory tasks (e.g., see Conway, 1996; Hunt, 2006; Rubin & Schulkind, 1997). Distinctiveness effects have been observed across a diverse array of manipulations, including semantic distinctiveness of individual words (e.g., Hunt & Mitchell, 1982; Schmidt, 1985), orthographic distinctiveness of verbal items (e.g., Hunt & Elliot, 1980), visual distinctiveness of facial features (e.g., Winograd, 1981), emotional distinctiveness of flashbulb memories (Brown & Kulik, 1977) or traumatic events (Loftus & Burns, 1982), olfactory distinctiveness for odor-evoked memories (e.g., Herz, 1997; Herz & Cupchik, 1995), memory for humorous text (Kintsch & Bates, 1977), and

---

Mark L. Howe, Lancaster University, Lancaster, United Kingdom.

Portions of this research were conducted when Mark L. Howe was at Lakehead University, Thunder Bay, Ontario, Canada, and were funded by Natural Sciences and Engineering Research Council of Canada Grant OGP0003334 to Mark L. Howe.

Correspondence concerning this article should be addressed to Mark L. Howe, Department of Psychology, Fylde College, Lancaster University, Lancaster LA1 4YF, England. E-mail: mark.howe@lancaster.ac.uk

bizarre imagery effects (e.g., Cox & Wollen, 1981; Einstein, McDaniel, & Lackey, 1989; Fritsch & Larsen, 1990; O'Brien & Wolford, 1982). Moreover, distinctiveness effects in adults' memory can have important forensic consequences. For example, distinctiveness has been shown to be related to the superior retention of traumatic events (for a review, see Howe, 1998a) and to reduce false memories (e.g., Arndt & Reder, 2003; Howe, 1998b; Smith & Hunt, 1998).

Although much is known about distinctiveness effects in adults' memory, the same cannot be said for children's memory. Admittedly, there is considerable developmental continuity in processes that govern storage and retrieval of information in children's and adults' memory, so much so that it frequently has been assumed that the effects of memory variables (e.g., distinctiveness) would be the same regardless of age (see Howe, 2000). Unfortunately, although it is true that there are numerous similarities in the fundamental mechanisms that regulate memory across age, it has not always been the case that factors affecting memory have the same effects regardless of age (see Howe, 2006). Indeed, as it turns out, distinctiveness may be one of those factors that exhibit not only quantitative variation with age but also qualitative variation (for a recent review, see Howe, in press). For example, Howe, Courage, Vernescu, and Hunt (2000) found that the isolation (or von Restorff) effect improved 5- and 7-year-olds' retention when the isolate differed from the otherwise categorized list on a perceptual dimension (color), but only the older children's performance improved when the isolate differed on a semantic dimension (different category exemplar or a numeral). In fact, for 5-year-olds, when number served as the isolate, performance significantly declined.

Other distinctiveness manipulations have produced similar effects. For example, although Howe et al. (2000) found no developmental differences in the magnitude of improvement in 5- and 7-year-olds' retention in a paired-associate task using bizarre imagery, others have found that only older (12- and 13-year-olds) children benefit from bizarre interactive imagery. As well, although Hudson (1988; also see Davidson & Hoe, 1993; Davidson & Jergovic, 1996) found that script-violating information was better remembered than script-consistent information, Farrar and Goodman (1992) did not find such a trend with children. Indeed, Farrar and Boyer-Pennington (1999) found that the degree to which distinctive information was better remembered than script-consistent information varied as a function of age (7-year-olds remembered better than 4-year-olds) as well as experience with the script materials (with more experience being better). Thus, it is clear that distinctiveness, depending on how it is manipulated, sometimes exhibits similar beneficial effects across age and at other times differs quantitatively and qualitatively across development in childhood.

More recently, distinctiveness effects have been examined in some rather nontraditional paradigms (see Howe, 1998b). For example, Ghetti, Qin, and Goodman (2002) examined the role of distinctiveness in reducing children's false memories using the Deese–Roediger–McDermott paradigm (Deese, 1959; Roediger & McDermott, 1995). The authors manipulated distinctiveness by presenting pictures of the concepts as well as presenting them orally. The results were very clear—the presentation of pictures along with words reduced false memory production in all children and, moreover, reversed the usual developmental finding whereby older children typically exhibit more false memories than younger children (e.g., Brainerd, Reyna, & Forrest, 2002; Howe, Cicchetti, Toth, & Cerrito, 2004). Thus, pictures at encoding provided a distinctive cue children could use when recalling what was on the list versus what was not on the list but implied. This study, along with the ones already discussed, shows the importance of distinctiveness when manipulated at the time information is being encoded and stored.

Only two recent studies have examined distinctiveness effects after the information has been encoded and stored. In a series of experiments, children's retroactive interference was shown to decline when the interfering material was made distinct in their memory (Howe, 2002, 2004). In these experiments, children learned two sets of information (either two stories or two lists of concepts) consecutively and 24 hr later were required to recall the first information set. In one experiment, the investigators made interfering material distinct by telling children to forget the material presented second (Howe, 2002). Although the to-be-forgotten material was not actually forgotten, the reduced levels of retroactive interference occurred because intentional forgetting instructions made the to-be-forgotten material distinct from the to-be-remembered material. In another experiment, children learned two categorized lists and were required to recode the second list of concepts using a new category (Howe, 2004). As with the intentional forgetting manipulation, children who recoded the second list experienced significantly reduced levels of retroactive interference because recoding made the interfering material distinctive in memory.

Several critical gaps remain in our understanding of distinctiveness effects in children's memory. First, because distinctiveness processing occurs, by definition (e.g., see Schmidt, 1991), against a backdrop of similarity, it is not known whether distinctiveness is any more effective in enhancing children's memory than additional similarity processing. That is, the effects of distinctiveness and similarity have never been directly compared in children's memory. Moreover, it is not known whether distinctiveness and similarity processing have similar effects on children's acquisition of information as they do on children's subsequent long-term retention of that information. Although some studies (with adults) have examined distinctiveness effects across retention intervals, distinctiveness effects at acquisition were confounded with differences at long-term retention (see Howe et al., 2000). This confound is avoided in the present study with controls for differences at acquisition (e.g., the use of criterion learning) and removal of their effects at long-term retention (e.g., via analysis of covariance [ANCOVA]).

The first purpose of the current experiment, then, is to extend the examination of distinctiveness effects in children's memory by directly examining the use of similarities and differences at acquisition and long-term retention. Similarities and differences were examined because both play an important role in learning and memory. Indeed, similarity (e.g., associative, semantic) aids the acquisition and retention of information, as seen, for example, when adults' and children's memory performance is better when the to-be-learned material is categorized than when it is unrelated (see Bjorklund, 2004; Howe, 2000). Differences also facilitate the acquisition and retention of information, particularly when that to-be-learned information is highly similar. Indeed, distinctiveness, no matter how it is defined, emerges only from a background of similarity. Again, to use the von Restorff procedure to elucidate, the isolate (e.g., dog) is different from the background category (e.g., vehicles: bicycle, car, train, jet) that constitutes the remaining items (or nonisolates) on the list. More generally, unique or distinctive processing, by definition, must occur against a background of highly organized similarity structures (also see Hunt & Smith, 1996). As Hunt (2006) said, "I suggest that this is exactly what we mean by the term distinctiveness: The processing of differences in the context of similarity" (p. 12).

Unlike the effects of similarity, which tend to facilitate trace acquisition and retention by increasing the internal cohesion of elements within a memory trace, distinctiveness can have equally beneficial effects on memory by reducing the probability of interference among like traces. However, it is not clear whether similarity and distinctiveness effects are due to changes in storage, retrieval, or both. Theoretical proposals have been advanced that favor storage processes or retrieval processes. For example, distinctiveness may operate at the time of encoding by enhancing the quantity and quality of the features sampled and stored from the target event. Alternatively, distinctiveness may serve to facilitate editing and censuring of information at the time of retrieval by "marking" information against the background noise of other memory information (for reviews, see Howe, 2006; Hunt, 2006). Thus, although it has been thought that distinctiveness benefits memory performance because it exerts its effect at retrieval by facilitating discriminative processes (e.g., Smith & Hunt, 1998), storage may also be involved inasmuch as better discriminated traces may be better discriminated because they stand out against the background noise of other, highly similar traces in storage. If this is true, then distinctiveness may affect retrieval through changes in storage. Developmentally, this is important because improvements in children's memory performance manifest themselves at both storage and retrieval, but when it comes to long-term retention, the evidence indicates that these improvements are due mainly to changes in storage (see Howe, 2000).

It is also not known whether the source of distinctiveness makes a difference in children's memory. That is, are distinctiveness effects more prominent when the distinctive information is generated by the children themselves or when it is generated by someone else? The second purpose of this experiment, then, is to examine the locus of children's self- versus other-generated elaboration effects in memory. Although it is known that children and adults remember more information when they generate their own elaborations than when experimenters generate them (e.g., see Bjorklund, 2004), it is not known whether these effects extend to distinctive processing. It is also not known whether these effects extend past acquisition to children's long-term retention. Moreover, like the effects of distinctiveness itself, it is not known whether any advantage of self- versus other-generated distinctiveness is related to storage processes, retrieval processes, or both (for a recent review of generation effects with adults, see Mulligan & Lozito, 2004). The prevailing view is that because the person himself or herself is creating the elaboration, any advantage in memory performance should be due primarily to storage factors (e.g., more elaborate trace, better integrated features in storage). That is, when the person generates his or her own elaboration, such an elaboration involves the person's own unique semantic organization and thus would be better tuned to the idiosyncratic aspects of that person's memory, making it more durable and hence more memorable. These generation effects might be more prominent in older than in younger children. This is because older children not only have more experience with their own memory but, more important, have a richer, more fully developed and organized knowledge base (for a review, see Bjorklund, 1987).

A third purpose of this research, then, is to determine whether storage or retrieval explanations (or some combination of both) of distinctiveness and elaboration effects are most appropriate. In accordance with this purpose, a formal method is adopted in which the rules that map these theoretical processes onto performance outcomes are made explicit (see Batchelder & Riefer, 1999). One model that has been used successfully in studying the contribution of storage and retrieval processes to the development of children's long-term retention processes is the trace-integrity model. Because lengthy reviews of this model and its mathematical and statistical properties have appeared a number of times in the recent literature (see Howe, 2000, 2002), only a summary of its key features are provided here (see also Appendix A for a more detailed presentation).

Trace-Integrity Theory

In the trace-integrity theory, storage and retrieval are viewed as processes lying on a continuum on which memory traces consist of collections of primitive elements (e.g., features, nodes). Learning consists of sampling elements from the to-be-remembered event and, along with elements already in memory, integrating these elements into a single, cohesive, and active structure in memory. Retention of information is determined by the degree to which that structure remains integrated and cohesive over time. Traces that are poorly integrated tend to disintegrate, and their stability (in terms of both availability—i.e., storage—and accessibility—i.e., retrieval) is compromised. When this occurs, the trace becomes unstable in memory, loses its distinctiveness, and fades into the background noise of other memory traces (see Howe, 2000). This view of how memory traces are formed and retained is consistent not only with other contemporary views of memory (e.g., Estes, 1988, 1997) but also with many current views of memory development (e.g., Bauer, 2004; Bauer, Wenner, Dropik, & Wewerka, 2000; Rovee-Collier, Hayne, & Colombo, 2001; Schneider & Bjorklund, 1998). A key advantage of the trace-integrity theory is that it comes with an associated mathematical model, one that permits the extraction of these proposed theoretical processes (storage and retrieval) from empirical outcomes on memory tests (test trial errors and successes).

Like other developmental theories, the trace-integrity theory includes a mechanism whereby memory improvements that occur with age are accounted for either directly, in terms of changes in basic memory processes (e.g., encoding, storage, retrieval), or indirectly, in terms of changes in other cognitive factors that are correlated with age (e.g., knowledge, strategies). According to trace-integrity theory, the acquisition of memory representations becomes easier because (a) trace integration improves as children's ability to organize features into cohesive patterns develops (e.g., through improved categorization and feature extraction) and (b) children's ability to create traces that stand out against the background noise of other memories improves as their knowledge base develops (e.g., creating traces that are distinctive; see Howe, 2000). Trace retention also depends on these two factors. That is, developmental advances in memory more generally, and retention more particularly, depend on corresponding changes in a host of cognitive factors that facilitate the cohesion of primitive trace elements or bundles of features in memory (e.g., knowledge;

distinctiveness; strategies, e.g., rehearsal; categorization; scripts), reduce the likelihood of their modification over time (e.g., through blending, recoding, reconstruction, reorganization) or disintegration (e.g., interference, trace decay), and facilitate trace redintegration (e.g., reinstatement, testing, retrieval cuing; also see Bjorklund, 1987; Howe, 2000; Schneider & Bjorklund, 1998).

Considerable evidence consistent with this theory has been reviewed elsewhere (e.g., Courage & Howe, 2004; Howe, 2000; Howe, Courage, & Edison, 2003). For present purposes, the major advantage of this theoretical approach is that it can help to decipher the storage and retrieval loci of distinctiveness and generation effects at long-term retention. To do this, a formal model of long-term retention is used to partition the storage and retrieval components of children's retention performance. Because this model applies to the long-term retention of information, participants in these studies are given more than a single study–test sequence on the to-be-remembered material. Children learn the information to a strict acquisition criterion of two consecutive errorless recall trials. There are three reasons for adhering to such a strict criterion. First, this avoids item difficulty confounds. That is, items that are easier to learn are not better represented in memory than more difficult items by the time criterion is achieved. Second, to the extent that distinctiveness effects may take time to emerge (especially with children), more than a single study–test opportunity may be necessary to observe such effects at initial acquisition. This is particularly relevant in studies in which children are presented with conceptual information, as categorical relations may be activated more slowly than in adults (e.g., see Bjorklund, 1988). As well, if distinctiveness is being measured at retention, failures to equate item differences at the end of acquisition will be confounded with item differences at retention. Third, when these item difficulty confounds are left uncontrolled in developmental research, age differences due to manipulations at acquisition will be confounded with age differences in forgetting rates (see comments in the Results section as well as Howe & Courage, 1997). Thus, in all of the experiments using the trace-integrity model, children learned the list information to a criterion of two consecutive errorless trials, which ensured both that (a) all of the information was stored in memory and was correctly retrieved at least twice in a row and (b) differences (itemwise as well as developmental) at retention could be evaluated independently of differences in initial learning difficulty.

The basic assumption of this mathematical model, like the theoretical model, is that storage and retrieval processes lie on a continuum, with the integrity of the trace varying with the extent to which trace elements (features, nodes, etc.) are bound together into a whole. As traces are learned, their elements are encoded and stored in an increasingly stable structure. Following acquisition, traces can either remain in the same stable state, lose some of their integrity (which compromises their retrievability), or undergo disintegration of their structural properties in storage over a retention interval. On subsequent tests of long-term retention, traces that do not lose their integrity exhibit correct response rates of 1.00 (no failures in storage or retrieval). Those that experience some loss of integrity (which makes them more difficult to access or retrieve) exhibit correct response probabilities that are less than 1.00 but greater than zero (retrieval failure only). Finally, those whose structural integrity is severely compromised exhibit correct response probabilities of zero (failures in both storage and retrieval).

In addition to these forgetting processes, there exist reminiscence probabilities both in storage and in retrieval. That is, even though a retention session may consist of test trials only, with no additional, explicit learning opportunities (i.e., re-presentation of list items), there exists the possibility that traces can be reminisced through reinstatement of retrieval processes (for those items that underwent retrieval-based forgetting) or redintegrated in storage (for those items that underwent both storage- and retrieval-based forgetting).

These basic ideas have been implemented in a mathematical model that pertains to experiments that use four test opportunities with no further study. A description of the various parameters that measure storage failures (S), retrieval failures (R), storage reminiscence (a), and retrieval reminiscence (r_i and f_i) is given in Table 1 (specifics on how these different parameters are mapped onto the relevant data space are provided in a more extensive appendix in Howe, 2000).

In what follows, this trace-integrity model (along with more traditional analyses) is used to examine the role of storage and retrieval processes in young children's (5- and 7-year-olds) 3-week retention of self- and experimenter-elaborated similar and distinctive materials. These two age groups were selected because they span a critical period in cognitive development, namely, the *5–7-year shift* (see Sameroff & Haith, 1996). It is during this time frame that children make the transition from using complementary to using conceptual classification (see Bjorklund, 2004). Such advances in children's understanding of conceptual relations might have predictable effects on their use of similarity and distinctive-

Table 1
Definitions of the Parameters in the Trace-Integrity Framework

Process and parameter	Description
Trace forgetting	
S	The probability of storage failure
R	The probability of retrieval failure of information in storage
Trace reminiscence	
a	The probability that information not in storage is redintegrated to a level above zero recall
r_1	The probability of two consecutive successes
r_2	The probability of three consecutive successes
r_3	The probability of four consecutive successes
f_1	The probability of a success following one error
f_2	The probability of a success following two consecutive errors
f_3	The probability of a success following three consecutive errors

ness, potentially enhancing their encoding, storage, and retrieval of information.

Method

Participants

The participants were 112 (56 male, 56 female) 5-year-olds (mean age = 5 years 4 months; range = 5 years 0 months to 5 years 11 months) and 110 (55 male, 55 female) 7-year-olds (mean age = 7 years 5 months; range = 7 years 0 months to 7 years 11 months). Ninety-five percent of the children were middle-class, were White, and came from a predominantly metropolitan area; 5% were Aboriginal. The children, all of who participated with written parental consent and their own assent, were randomly selected from classes in their schools.

Materials

Children learned a list of 18 related paired associates (see Appendix B). Each pairing was depicted with black-and-white pictures that were drawn from the most frequent category exemplars, representing 18 different categories (Battig & Montague, 1969; Posnansky, 1978). All of the materials were rated high in frequency, familiarity, concreteness, and imageability for both adults and children (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997; Paivio, Yuille, & Madigan, 1968).

Design and Procedure

The design was a simple 2 (age: 5-year-olds vs. 7-year-olds) × 5 (condition: control vs. experimenter-generated similarities vs. experimenter-generated differences vs. self-generated similarities vs. self-generated differences) × 4 (test trials: 1–4 at long-term retention) experiment. All of the children learned the same list (to a criterion of two consecutive errorless test trials) and returned 3 weeks later for a four-trial long-term retention test on that list. In the control condition, children were instructed to learn the 18 pairs given standard paired-associate learning instructions. In addition, for each pair, the child was given the category label (e.g., "Here are two animals, a dog and a cat"). This was done to be consistent with the four instructed conditions, in which children were given the category label and either provided with an additional similarity (e.g., "Both dogs and cats have four legs"; experimenter-generated similarities), asked to provide a similarity themselves (self-generated similarities), provided with a difference between the two animals (e.g., "Dogs bark; cats meow"; experimenter-generated differences), or asked to provide a difference themselves (self-generated differences). Category labels were provided to all of the children to avoid potential differences in the spontaneous use of category information (something that is more likely in older than in younger children) and to prevent more abstract (category) information serving as the basis for similarity judgments. Rather, the intent was for all of the similarities and differences to be perceptual and concrete (lists of experimenter-generated items and samples of the most frequent child self-generated similarities and differences are provided in Appendix B).

Children were tested individually. The picture pairs were presented and named by the experimenter at a 5-s pace. Following this, children were tested on each of the pairs in a random cuing order. To prevent short-term memory effects, the first few cues did not come from the last few presented pairs at study, and the first pairs studied on the subsequent study trial were not among those tested last. This study–test procedure continued until the child was able to recall all of the list members correctly on two consecutive test trials. Across trials, items at study and test were presented in a quasi-random fashion to avoid serial position effects.

Three weeks following acquisition, children were administered four test trials on the list they had learned, without additional study opportunities. This repeated-trials procedure provides greater sensitivity in retrieving information from memory than the more usual single-trial procedure, possibly through the honing of retrieval skills, which make items that are available in memory more accessible through reminiscence processes. As at acquisition, cues were presented in a quasi-random order to avoid serial position and short-term memory effects.

Results

Three sets of findings are reported—an analysis of variance (ANOVA) was used to analyze the acquisition data, an ANCOVA was used to analyze the retention data, and the application of the trace-integrity model was used to ascertain the storage or retrieval locus of distinctiveness effects. Note that the ANCOVA was selected because it is one means by which researchers can control for possible under- and overlearning effects at long-term retention resulting from individual differences that carry over from initial acquisition. Although overlearning effects are generally rare (see Brainerd & Reyna, 1990, 1995; Howe & Courage, 1997), underlearning effects may be more troublesome. This is because criterion performance designs almost always require that slower learners (younger children) receive more study–test sequences (learning opportunities) to reach the same level of performance as the faster learners (older children). What this means is that in performance-based criterion-learning designs, additional learning trials for younger children might compensate for faster forgetting observed in slower (younger) learners, and this could obscure true differences in age-based forgetting rates (see Howe & Courage, 1997). When performance differences during acquisition are used as a covariate (average number correct at acquisition), any confounds that remain following (or might be caused by) criterion acquisition can be eliminated from the dependent variable (number of correct responses per trial at long-term retention) before the impact of the manipulations of interest is assessed.

Acquisition Data

The number of correct responses on each trial at acquisition was analyzed via a 2 (age) × 5 (condition) × 6 (acquisition trials) ANOVA. The results indicated a main effect for age, $F(1, 212) = 26.07$, $p < .001$ ($\eta^2 = .11$), whereby 7-year-olds ($M = 17.60$) had more average correct responses across the six acquisition trials than 5-year-olds ($M = 17.06$). There was also a main effect for condition, $F(4, 212) = 16.22$, $p < .001$ ($\eta^2 = .23$), whereby, according to post hoc (Neuman–Keuls) tests ($p < .05$), the average number of correct responses across all trials was lower in the control condition ($M = 16.77$) than in the two experimenter-generated conditions (similarities, $M = 17.07$; differences, $M = 17.23$), which did not differ; the control condition and the two experimenter-generated conditions, in turn, were lower than the two self-generated conditions (similarities, $M = 17.70$; differences, $M = 17.86$), which did not differ. Finally, there was also a main effect for trial, $F(5, 1060) = 220.91$, $p < .001$ ($\eta^2 = .51$), whereby the average number of correct responses increased over trials. There were two first-order interactions, Trial × Age, $F(5, 1060) = 13.75$, $p < .001$, and Trial × Condition, $F(20, 1060) = 17.62$, $p < .001$, both of which were modified by a second-order Trial × Age × Condition interaction, $F(20, 1060) = 7.11$, $p < .002$ ($\eta^2 = .04$), the results of which are depicted in Figure 1 (Panel A = 5-year-olds; Panel B = 7-year-olds). As can be seen from this figure and was confirmed by post hoc tests (Neuman–Keuls, $p < .05$), the interaction was obtained because (a) age

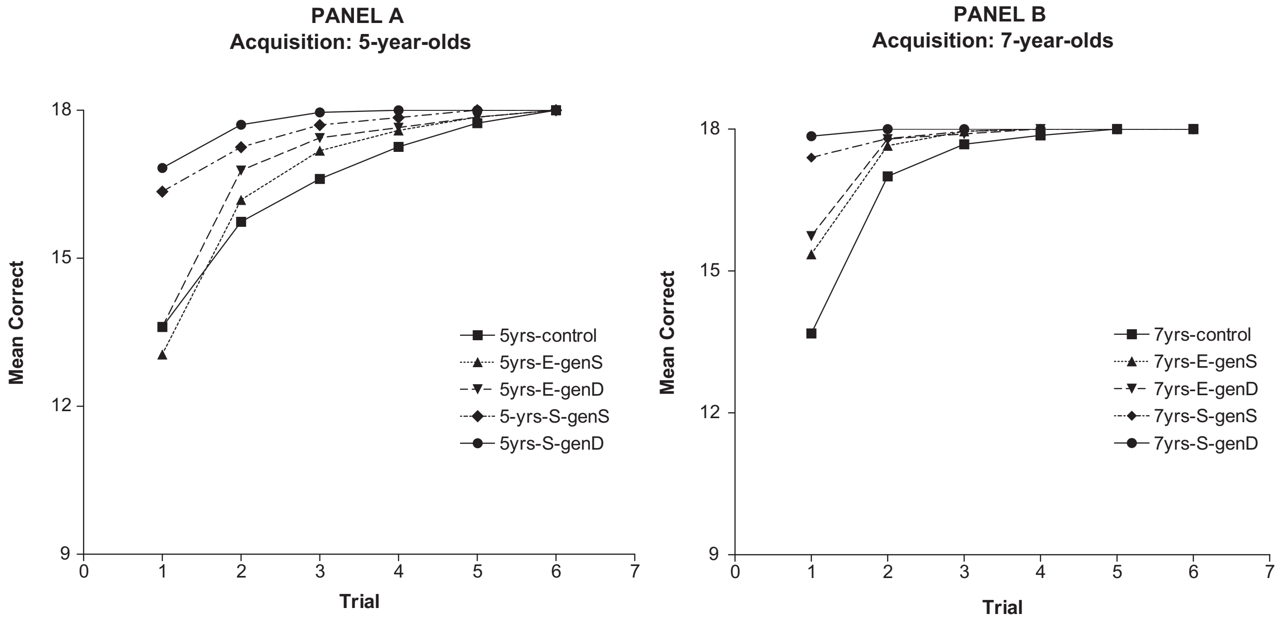

Figure 1. Mean number of responses correct at acquisition as a function of condition and trial for 5-year-olds (Panel A) and for 7-year-olds (Panel B). yrs = year-olds; E-genS = experimenter-generated similarities; E-genD = experimenter-generated differences; S-genS = self-generated similarities; S-genD = self-generated differences.

differences were present in the initial (intercept) starting values such that for 5-year-olds, initial performance was poorest for the control and experimenter-generated lists relative to the self-generated lists, whereas for 7-year-olds, there was a trichotomy (control < experimenter-generated lists < self-generated lists); (b) criterion performance was achieved more quickly by older than by younger children; (c) criterion performance was achieved more quickly on self- than on experimenter-generated lists, regardless of age; and (d) criterion performance was achieved more quickly on lists that involved differences than on lists that involved similarities. Overall, then, older children learned the lists faster than younger children; both similarities and differences helped children learn the paired associates, with self-generated items being more quickly learned than experimenter-generated items; and differences, regardless of who generated them, were a better aid to learning than similarities.

Retention Data

Concerning retention, the average number of correct responses at acquisition served as the covariate, and the number of correct responses on each of the four trials was the dependent variable. These data were analyzed with a 2 (age) × 5 (condition) × 4 (retention trials) ANCOVA. The results indicated a significant effect of the covariate, $F(1, 211) = 5.66$, $p < .02$ ($\eta^2 = .026$). As well, there was a main effect for age, $F(1, 212) = 9.43$, $p < .002$ ($\eta^2 = .043$), whereby the average number correct was higher for 7-year-olds ($M = 15.65$) than for 5-year-olds ($M = 14.89$). There was also a main effect for condition, $F(4, 212) = 3.68$, $p < .007$ ($\eta^2 = .065$), whereby, according to post hoc (Neuman–Keuls) tests ($p < .05$), average correct recall exhibited the following pattern: The control condition ($M = 14.94$) was equal to both experimenter-generated conditions (similarities, $M = 14.98$; differences, $M = 14.72$), which did not differ from each other but were less than both self-generated conditions (similarities, $M = 15.92$; differences, $M = 15.78$), which did not differ from each other. Finally, there was an effect of trial, $F(3, 636) = 51.41$, $p < .001$ ($\eta^2 = .195$), whereby, according to post hoc (Neuman–Keuls) tests ($p < .05$), the average number correct increased across the four test trials, with significant differences occurring only between the first test trial and the remaining three. These data (plotted as a function of age and condition) can be seen in Figure 2 (Panel A = 5-year-olds; Panel B = 7-year-olds).

Overall, both differences and similarities were important for retention, with the main factor affecting retention being whether participants generated these similarities and differences or someone else did. As shown in the next section, according to a more sensitive model of retention performance, similarities and differences were important in children's retention, and these effects were distinguishable from those due to whether the experimenter or the child generated them.

Trace-Integrity Analyses

As already noted, the statistical and mathematical mechanics of this model are provided in Appendix A and are available in a number of prior sources (e.g., Howe, 2000, 2002) for the interested reader. Before the trace-integrity model can be used to interpret the impact of distinctiveness manipulations on children's long-term retention, the goodness of fit of the model must be evaluated. That is, it must be shown that the model, with only 9 parameters, gives as full an accounting of the data as a purely empirical model in which all 15 degrees of freedom (empirical parameters) are free to vary in the data. The likelihood ratio tests (one was conducted for

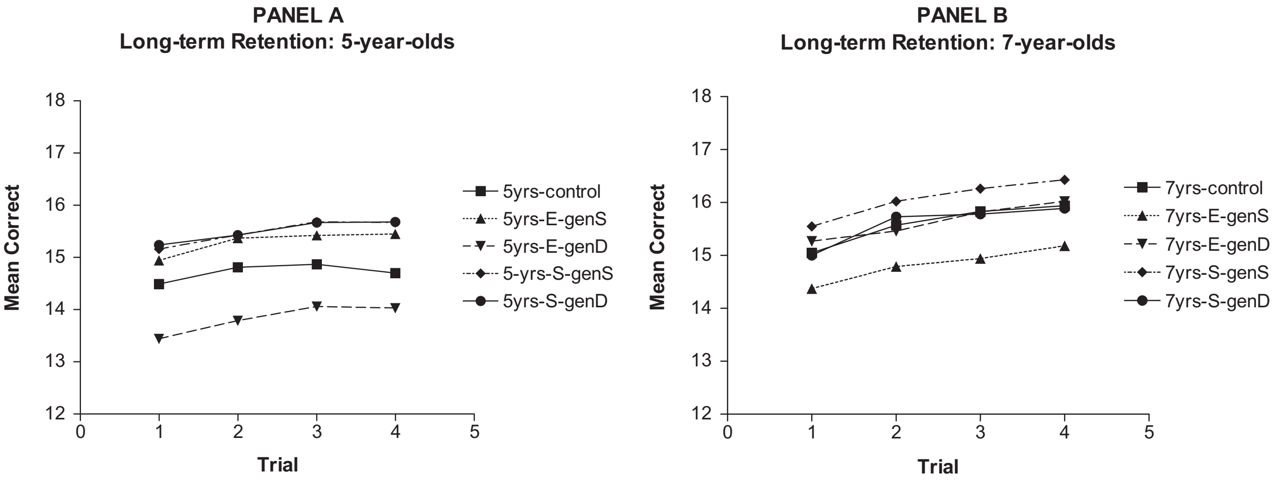

Figure 2. Mean number of responses correct at retention as a function of condition and trial for 5-year-olds (Panel A) and for 7-year-olds (Panel B). yrs = year-olds; E-genS = experimenter-generated similarities; E-genD = experimenter-generated differences; S-genS = self-generated similarities; S-genD = self-generated differences.

each condition; tests contained 6 degrees of freedom) evaluate the null hypothesis that the theoretical model does as good a job of accounting for the data as the empirical model. The results of the 10 tests for this experiment failed to reject the null hypothesis, indicating that the model provided an adequate fit to the retention data generated across the conditions, all χ^2s(6) < 10.26.

The numerical values of the parameters for this experiment are shown in Table 2. To see whether the values of the theoretical parameters of the trace-integrity model differed significantly across the various conditions in this experiment, an experiment-wise test was conducted. Like the omnibus *F* test, this experiment-wise test evaluated the null hypothesis that, on average, the model's parameters did not vary among conditions. For the data in the present experiment, the value of this test statistic was $\chi^2(81) = 160.20$, $p < .01$.

The next step involved a series of condition-wise tests, which evaluated the null hypothesis that the numerical values of the parameters did not vary between specific pairs of conditions. In this experiment, a total of 21 condition-wise tests were needed, 5 comparisons to evaluate age effects and 8 within each age to evaluate distinctiveness and generation effects. The numerical results of the tests to evaluate age effects (all *p*s $< .01$) revealed differences between 5-year-olds and 7-year-olds on all lists, all χ^2s(9) > 21.67. Two sets of comparisons were conducted to evaluate distinctiveness effects, one set between the control condition and the other four conditions at each age and the second set

Table 2
Parameter Estimates

Age and condition	S	R	a	r_1	r_2	r_3	f_1	f_2	f_3
				5-year-olds					
Control	.23	.16	.04	.97	1.00	.97	.33	.00	.00
E-generated similarities	.23	.16	.08	.99	1.00	.99	.30	.03	.06
E-generated differences	.22	.05	.05	.98	.99	.99	.45	.00	.00
S-generated similarities	.12	.14	.03	.99	.99	1.00	.40	.03	.00
S-generated differences	.12	.02	.02	1.00	1.00	1.00	.34	.00	.00
				7-year-olds					
Control	.13	.12	.06	.98	1.00	1.00	.48	.06	.02
E-generated similarities	.11	.12	.00	.99	1.00	1.00	.35	.07	.00
E-generated differences	.12	.02	.05	.98	.99	1.00	.50	.00	.00
S-generated similarities	.02	.11	.07	1.00	1.00	1.00	.00	.09	.09
S-generated differences	.02	.01	.03	1.00	1.00	1.00	.00	.08	.05

Note. S = the probability of storage failure; R = the probability of retrieval failure of information in storage; a = the probability that information not in storage is redintegrated to a level above zero recall; r_1, r_2, and r_3 = the probability of two, three, or four consecutive successes, respectively; f_1, f_2, and f_3 = the probability of a success following one, two, or three consecutive errors, respectively; E-generated = experimenter-generated; S-generated = self-generated.

between the generation of similarities and the generation of differences at each age. Concerning the first set, the numerical results of the tests evaluating distinctiveness effects indicated that the control condition differed from all of the other conditions—control versus experimenter-generated differences, χ^2s(9) = 24.13 (5-year-olds) and 25.22 (7-year-olds); control versus self-generated similarities, χ^2s(9) = 25.73 (5-year-olds) and 28.72 (7-year-olds); control versus self-generated differences, χ^2s(9) = 37.83 (5-year-olds) and 24.72 (7-year-olds)—except the experimenter-generated similarities, χ^2s(9) = 6.30 (5-year-olds) and 9.07 (7-year-olds). Concerning the second set, the numerical results of the tests evaluating distinctiveness effects indicated that the similarities conditions differed from the differences conditions for both the 5-year-olds, χ^2s(9) = 22.18 (experimenter-generated) and 29.80 (self-generated), and the 7-year-olds, χ^2s(9) = 23.05 (experimenter-generated) and 26.57 (self-generated). Finally, a series of tests was conducted to evaluate the effects of who (experimenter or participant) generated the similarities and differences. The numerical results of the tests evaluating these generation effects indicated that the experimenter-generated conditions differed from the self-generated conditions for both the 5-year-olds, χ^2s(9) = 28.75 (similarities) and 36.31 (differences), and the 7-year-olds, χ^2s(9) = 22.34 (similarities) and 29.59 (differences).

The last sequence of tests, namely, parameter-wise tests, are used to isolate the specific parameter or parameters whose estimates differed between the conditions. Each of these chi-square tests has one degree of freedom, and, because they are tedious and space consuming to report, they are often given in summary form. Consistent with this tradition, only those parameter-wise differences that were significant ($p < .05$) are reported next. As well, because age, distinctiveness, and generation effects were localized primarily in terms of forgetting, not reminiscence, rates, a phenomenon that is typical in this area (see Howe, 2000, 2002, 2004), the discussion of the model-based findings is confined to the parameters that measure storage failures (S) and retrieval failures (R).

Like the previous analyses, there are three issues to focus on here, namely, the effects of age, generation, and distinctiveness. Concerning age, the effects obtained were typical for long-term retention research (e.g., Howe, 2000, 2002, 2004; Howe et al., 2000)—that is, older children were better than younger children at maintaining traces in storage; the average difference between ages on the parameter S was .10. This is consistent with Howe's (2000) contention that much of what develops in children's long-term retention is the ability to maintain information in storage.

Concerning experimenter- versus self-generated similarities and differences, it is clear that these effects, too, were localized at storage—that is, regardless of age, children exhibited better storage-based retention of self- than of experimenter-generated similarities and differences (S mean difference = .11 for 5-year-olds and .10 for 7-year-olds). These findings are consistent with the speculation that self-generated elaborations benefit memory storage through either increased numbers of features associated with the memory trace, greater cohesion among the elements within the trace, or both.

Finally, as predicted, distinctiveness effects were localized at retrieval. That is, regardless of age, children exhibited better retrieval at retention for distinctive than for similar pairings, and this effect was the same regardless of age and whether differences and similarities were experimenter-generated (R mean difference = .11 for 5-year-olds and .10 for 7-year-olds) or self-generated (R mean difference = .12 for 5-year-olds and .10 for 7-year-olds).

Discussion

Three questions were examined in this research. First, does children's retention of related pairs benefit from additional similarity or distinctiveness processing? Second, does it matter who generates these similarities or differences, the children themselves or the experimenters? Third, are these effects related to storage, retrieval, or both? The discussion of this last question is integrated into the discourse associated with the first two questions.

Concerning the first question, the answer appears to be that both similarity and difference processing enhance children's acquisition and long-term retention of related concepts. That is, relative to a control condition, children's acquisition and long-term retention were enhanced when additional similarities or differences were generated between pair members. Moreover, at acquisition, criterion learning was achieved more rapidly on lists involving differences than on those involving similarities. At retention, traditional analyses did not reveal differences between similarity and distinctiveness processing, although both were better than the control condition. However, the trace-integrity model did reveal a very important difference. That is, items that underwent additional distinctive rather than similarity processing were easier to retrieve. Thus, like the findings with adults (e.g., Smith & Hunt, 2000), the present results follow directly from the assumptions outlined earlier about the nature of distinctive processing. The fact that additional difference information about categorized material produced better learning and retention than additional similarity information (whether experimenter- or self-generated) can be understood as the beneficial effect of distinctiveness processing. Although additional similarity was beneficial for children's memory of related pairs, distinctiveness had a more powerful effect at both acquisition and long-term retention.

Concerning the second question, self-generated similarities and differences produced superior memory performance to experimenter-generated ones, both at acquisition and at retention. This general finding is consistent with previous research on children's elaborative memory processing (for a review, see Bjorklund, 2004). What the current findings add is that these effects extended to children's long-term retention and to the processing of both similarities and differences. Equally important, this is the first time it has been shown that these effects, at least at retention, are storage based. That is, regardless of age, self-generated information was better maintained in storage over a 3-week retention interval than experimenter-generated information. As noted earlier, this is consistent with the speculation that self-generated elaboration results in more and better integrated features sets that compose the memory trace.

It is interesting that, despite the typical age effects whereby older children learned more rapidly than younger children and retained more information over the retention interval, these differences were not modulated by distinctiveness or generation effects. Indeed, the magnitude of storage differences at retention, the primary source of age differences according to the trace-integrity model, did not vary as a function of self- versus experimenter-generated similarities or differences. This is similar to some previous research, in which the magnitude of distinctiveness effects did not vary across age (Howe et al., 2000). Thus, although there is reason to anticipate an interaction between age and distinctiveness (or generation effects), such interactions do not routinely turn up in studies examining children's long-term retention, especially when the memory task is relatively easy (as in this case, in which children learned conceptually related picture pairs given category labels). That is, distinctiveness and generation effects are developmentally invariant, at least in the age range studied in this article. This invariance notwithstanding, additional processing advances cannot be ruled out by the present data. As with other memory developments that tend to occur during later childhood and adolescence, the pattern of distinctiveness and generation effects at long-term retention may be altered by corresponding increases in the speed or automaticity of semantic access, changes that are correlated with developments in children's knowledge base (also see Howe, 2005).

More important, like previous research using the trace-integrity model, the main developmental difference in children's retention was a decreased susceptibility to storage failure (see also Bauer, 2005). Additional gains in storage competence were seen when children generated their own elaborations rather than simply using the ones provided by the experimenter. Unlike these storage-based retention effects, distinctiveness effects were localized at retrieval. As noted, this latter finding is consistent with much theoretical speculation in the adult literature on the locus of distinctiveness effects in memory (e.g., see Howe, 2006; Hunt, 2006; Smith & Hunt, 2000). What this experiment shows is that these findings extend to children's memory as well. Although distinctiveness effects can and do exist at the level of information storage (see Howe, 2002, 2004; Howe et al., 2000), learning paradigms such as the one used in this study routinely show the effects of distinctiveness at retrieval. When storage demands are reduced, as they are in the retention of paired associates, robust retrieval effects are anticipated, ones that prevent competition or interference of similar items at the time of output. Therefore, not only are distinctiveness effects themselves developmentally invariant in childhood and into adulthood (with the possible exception of changes in automaticity or speed of semantic access), but the locus of these effects is also invariant across age.

Overall, the findings from this experiment show that there are two effects that conspire to produce better acquisition and retention in young children. The first concerns factors that reduce storage failure rates. In the current experiment, having children generate their own similarities and differences between related conceptual pairs improved storage maintenance across a 3-week retention interval. This finding is consistent with the trace-integrity theory's claim that reduction of storage failure rates is an important component of children's memory development (see Howe, 2000; also see Bauer, 2005). The second effect concerns factors that enhance the retrievability of information. In the current experiment, generating differences rather than similarities between concepts within a pair facilitated the discriminative process at retrieval. That is, as with adults, distinctive processing aided children's ability to identify the correct target for output and eliminate related but incorrect items (Smith & Hunt, 2000). Clearly, in this instance, although older children were better at learning and remembering than younger children, like adults, children's storage benefited from the generation of additional elaborative information, and their retrieval was enhanced by distinctiveness. Thus, distinctiveness effects, at least in terms of their memory-enhancing properties and their locus (storage, retrieval) in basic memory processes, are developmentally invariant.

References

Arndt, J., & Reder, L. M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory and Language, 48,* 1–15.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86.

Battig, W. F., & Montague, W. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology, 80*(3, Pt. 2), 1–46.

Bauer, P. J. (2004). Getting explicit memory off the ground: Steps toward construction of a neuro-developmental account of changes in the first two years of life. *Developmental Review, 24,* 347–373.

Bauer, P. J. (2005). Developments in declarative memory: Decreasing susceptibility to storage failure over the second year of life. *Psychological Science, 16,* 41–47.

Bauer, P. J., Wenner, J. A., Dropik, P. L., & Wewerka, S. S. (2000). Parameters of remembering and forgetting in the transition from infancy to childhood. *Monographs of the Society for Research in Child Development, 65*(4, Serial No. 263).

Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: An interpretive review. *Developmental Review, 7,* 93–130.

Bjorklund, D. F. (1988). Acquiring a mnemonic: Age and category knowledge effects. *Journal of Experimental Child Psychology, 45,* 71–87.

Bjorklund, D. F. (2004). *Children's thinking: Development and individual differences* (4th ed.). Belmont, CA: Wadsworth/Thompson.

Brainerd, C. J., & Reyna, V. F. (1990). Can Age × Learnability interactions explain the development of forgetting? *Developmental Psychology, 26,* 194–203.

Brainerd, C. J., & Reyna, V. F. (1995). Learning rate, learning opportunities, and the development of forgetting. *Developmental Psychology, 31,* 251–262.

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are young children susceptible to the false memory illusion? *Child Development, 73,* 1363–1377.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5,* 73–99.

Conway, M. A. (1996). Autobiographical knowledge and autobiographical memories. In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 67–93). New York: Cambridge University Press.

Courage, M. L., & Howe, M. L. (2004). Advances in early memory development research: Insights about the dark side of the moon. *Developmental Review, 24,* 6–32.

Cox, S. D., & Wollen, K. A. (1981). Bizarreness and recall. *Bulletin of the Psychonomic Society, 18,* 244–245.

Csikszentmihalyi, M., & Beattie, O. V. (1979). Life themes: A theoretical and empirical exploration of their origins and effects. *Journal of Humanistic Psychology, 19,* 45–63.

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65,* 171–237.

Davidson, D., & Hoe, S. (1993). Children's recall and recognition memory for typical and atypical actions in script-based stories. *Journal of Experimental Child Psychology, 55,* 104–126.

Davidson, D., & Jergovic, D. (1996). Children's memory for atypical actions in script-based stories: An examination of the disruption effect. *Journal of Experimental Child Psychology, 61,* 134–152.

Deese, J. (1959). On the prediction of occurrence of certain verbal intrusions in free recall. *Journal of Experimental Psychology, 58,* 17–22.

Eacott, M. J., & Crawley, R. A. (1998). The offset of childhood amnesia:

Memory for events that occurred before age 3. *Journal of Experimental Psychology: General, 127,* 22–33.

Einstein, G. O., McDaniel, M. A., & Lackey, S. (1989). Bizarre imagery, interference, and distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 137–146.

Estes, W. K. (1988). Human learning and memory. In R. C. Anderson, R. J. Hernstein, G. Lindzey, & R. D. Luce (Eds.), *Steven's handbook of experimental psychology: Vol. 2. Learning and cognition* (2nd ed., pp. 351–415). New York: Wiley.

Estes, W. K. (1997). Processes of memory loss, recovery, and distortion. *Psychological Review, 104,* 148–169.

Farrar, M. J., & Boyer-Pennington, M. E. (1999). Remembering specific episodes of a scripted event. *Journal of Experimental Child Psychology, 73,* 266–288.

Farrar, M. J., & Goodman, G. S. (1992). Developmental changes in event memory. *Child Development, 63,* 173–187.

Fritsch, T., & Larsen, J. D. (1990). Image-formation time is not related to recall of bizarre and plausible images. *Perceptual and Motor Skills, 70,* 1259–1266.

Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology, 38,* 705–718.

Herz, R. S. (1997). The effects of cue distinctiveness on odor-based context-dependent memory. *Memory & Cognition, 25,* 375–380.

Herz, R. S., & Cupchik, G. C. (1995). The emotional distinctiveness of odor-evoked memories. *Chemical Senses, 20,* 517–528.

Howe, M. L. (1998a). Individual differences in factors that modulate storage and retrieval of traumatic memories. *Development and Psychopathology, 10,* 681–698.

Howe, M. L. (1998b). When distinctiveness fails, false memories prevail. *Journal of Experimental Child Psychology, 71,* 170–177.

Howe, M. L. (2000). *The fate of early memories: Developmental science and the retention of childhood experiences.* Washington, DC: American Psychological Association.

Howe, M. L. (2002). The role of intentional forgetting in reducing children's retroactive interference. *Developmental Psychology, 38,* 3–14.

Howe, M. L. (2004). The role of conceptual recoding in reducing children's retroactive interference. *Developmental Psychology, 40,* 131–139.

Howe, M. L. (2005). Children (but not adults) can inhibit false memories. *Psychological Science, 16,* 927–931.

Howe, M. L. (2006). The role of distinctiveness in children's memory. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 237–257). New York: Oxford University Press.

Howe, M. L., Cicchetti, D., Toth, S. L., & Cerrito, B. M. (2004). True and false memories in maltreated children. *Child Development, 75,* 1402–1417.

Howe, M. L., & Courage, M. L. (1997). Independent paths in the development of infant learning and forgetting. *Journal of Experimental Child Psychology, 67,* 131–163.

Howe, M. L., Courage, M. L., & Edison, S. C. (2003). When autobiographical memory begins. *Developmental Review, 23,* 471–494.

Howe, M. L., Courage, M. L., Vernescu, R., & Hunt, M. (2000). Distinctiveness effects in children's long-term retention. *Developmental Psychology, 36,* 778–792.

Hudson, J. A. (1988). Children's memory for atypical actions in script-based stories: Evidence for a disruption effect. *Journal of Experimental Child Psychology, 46,* 159–173.

Hunt, R. R. (2006). What is the meaning of distinctiveness for memory research? In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 3–25). New York: Oxford University Press.

Hunt, R. R., & Elliot, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General, 109,* 49–74.

Hunt, R. R., & Mitchell, D. B. (1982). Independent effects of semantic and nonsemantic distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 81–87.

Hunt, R. R., & Smith, R. E. (1996). Accessing the particular from the general: The power of distinctiveness in the context of organization. *Memory & Cognition, 24,* 217–225.

Kintsch, W., & Bates, E. (1977). Recognition memory for statements from a classroom lecture. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 150–159.

Loftus, E. F., & Burns, T. E. (1982). Mental shock can produce retrograde amnesia. *Memory & Cognition, 10,* 318–323.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254–278.

Mulligan, N. W., & Lozito, J. P. (2004). Self-generation and memory. *Psychology of Learning and Motivation, 45,* 175–214.

O'Brien, E. J., & Wolford, C. R. (1982). Effect of delay in testing on retention of plausible versus bizarre mental images. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 148–152.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76*(1, Pt. 2), 1–25.

Posnansky, C. J. (1978). Category norms for verbal items in 25 categories for children in grades 2–6. *Behavior Research Methods & Instrumentation, 10,* 819–832.

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented on lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 21, 803–814.

Rovee-Collier, C., Hayne, H., & Colombo, M. (2001). *The development of implicit and explicit memory.* Amsterdam: John Benjamins.

Rubin, D. C., & Schulkind, M. D. (1997). Distribution of important word-cued autobiographical memories in 20-, 35-, and 70-year-old adults. *Psychology and Aging, 12,* 524–535.

Sameroff, A. J., & Haith, M. (Eds.). (1996). *The five to seven year shift.* Chicago: University of Chicago Press.

Schmidt, S. R. (1985). Encoding and retrieval processes in the memory for conceptually distinctive events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 565–578.

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition, 19,* 523–542.

Schneider, W., & Bjorklund, D. F. (1998). Memory. In W. Damon (Series Ed.), D. Kuhn (Vol. Ed.), & R. S. Siegler (Vol. Ed.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th edition). New York: Wiley.

Smith, R. E., & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review, 5,* 710–715.

Smith, R. E., & Hunt, R. R. (2000). The effects of distinctiveness require reinstatement of organization: The importance of intentional memory instructions. *Journal of Memory and Language, 43,* 431–446.

Usher, J. A., & Neisser, U. (1993). Childhood amnesia and the beginnings of memory for four early life events. *Journal of Experimental Psychology: General, 122,* 155–165.

Winograd, E. (1981). Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 181–190.

Appendix A

Trace-Integrity Model

The parameters from Table 1 are mapped onto the outcome space of a four-trial retention experiment in Table A1. This mapping provides for the lawful extraction of information about storage and retrieval processes from observable outcomes on test trials. That is, this set of equations states explicitly what the relationship is between the outcomes of an experiment and the underlying or latent memory and cognitive processes that determine these outcomes (also see Batchelder & Riefer, 1999). A detailed description of how these parameters are linked to the data, the basis for the interpretation of the parameters, the history of the model, and the utility of this approach for the exploration of developmental differences in retention were provided in a recent review by Howe (2000).

The purpose of this appendix is to show the statistical machinery used to evaluate the model's goodness of fit to the observations and the procedures used to evaluate hypotheses about theoretical processes. First, it is paramount that before the trace-integrity model is used to interpret the impact of any manipulation on the theoretical processes measured in this framework, the goodness of fit of the model to the data must be evaluated and numerical estimates of the parameters must be obtained. Fortunately, a well-defined set of procedures is available for these purposes (see Howe, 2000, 2002). The process involves the five-step sequence outlined next.

The first two steps involve translating the data space into a probability space, which is subsequently transformed into a mathematical space. According to maximum likelihood theory, the first transformation results in a function that expresses the a posteriori likelihood of any data sample. The present function has 15 degrees of freedom and takes the form

$$\boldsymbol{L}_{15} = \{p[\mathrm{C}_1\mathrm{C}_2\mathrm{C}_3\mathrm{C}_4]^{\mathrm{N[CCCC]}} \times p[\mathrm{C}_1\mathrm{C}_2\mathrm{C}_3\mathrm{E}_4]^{\mathrm{N[CCCE]}} \times \cdots \times p[\mathrm{E}_1\mathrm{E}_2\mathrm{E}_3\mathrm{C}_4]^{\mathrm{N[EEEC]}} \times p[\mathrm{E}_1\mathrm{E}_2\mathrm{E}_3\mathrm{E}_4]^{\mathrm{N[EEEE]}}\}. \quad \text{(A1)}$$

The second transformation permits derivation of the theoretical likelihood, something that is accomplished by the simple substitution of the equations in Table A1 (denoted in Equation A2 by the term h) for the 16 terms in Equation A1. This yields a function with 9 degrees of freedom (because the 16 expressions are based on only 9 parameters) and takes the form

$$\boldsymbol{L}_9 = \{h(p[\mathrm{C}_1\mathrm{C}_2\mathrm{C}_3\mathrm{C}_4])^{\mathrm{N[CCCC]}} \times h(p[\mathrm{C}_1\mathrm{C}_2\mathrm{C}_3\mathrm{E}_4])^{\mathrm{N[CCCE]}} \times \cdots \times h(p[\mathrm{E}_1\mathrm{E}_2\mathrm{E}_3\mathrm{C}_4])^{\mathrm{N[EEEC]}} \times h(p[\mathrm{E}_1\mathrm{E}_2\mathrm{E}_3\mathrm{E}_4])^{\mathrm{N[EEEE]}}\}. \quad \text{(A2)}$$

The third step consists of counting the number of times each of the 16 events occurred in the sample data (i.e., summing across

Table A1
Mathematical Expressions Defining the Empirical Outcome Space

Outcome	Expression
p(CCCC)	$(1 - S)(1 - R)r_1r_2r_3$
p(CCCE)	$(1 - S)(1 - R)r_1r_2(1 - r_3)$
p(CCEC)	$(1 - S)(1 - R)r_1(1 - r_2)f_1$
p(CECC)	$(1 - S)(1 - R)(1 - r_1)f_1r_1$
p(ECCC)	$Sa(1 - R)r_1r_2 + (1 - S)Rf_1r_1r_2$
p(CCEE)	$(1 - S)(1 - R)r_1(1 - r_2)(1 - f_1)$
p(CECE)	$(1 - S)(1 - R)(1 - r_1)f_1(1 - r_1)$
p(ECCE)	$Sa(1 - R)r_1(1 - r_2) + (1 - S)Rf_1r_1(1 - r_2)$
p(CEEC)	$(1 - S)(1 - R)(1 - r_1)(1 - f_1)f_2$
p(ECEC)	$Sa(1 - R)(1 - r_1)f_1 + (1 - S)Rf_1(1 - r_1)f_1$
p(EECC)	$S(1 - a)a(1 - R)r_1 + SaRf_1r_1 + (1 - S)R(1 - f_1)f_2r_1$
p(CEEE)	$(1 - S)(1 - R)(1 - r_1)(1 - f_1)(1 - f_2)$
p(ECEE)	$Sa(1 - R)(1 - r_1)(1 - f_1) + (1 - S)Rf_1(1 - r_1)(1 - f_1)$
p(EECE)	$S(1 - a)a(1 - R)(1 - r_1) + SaRf_1(1 - r_1) + (1 - S)R(1 - f_1)f_2(1 - r_1)$
p(EEEC)	$S(1 - a)^2a(1 - R) + S(1 - a)aRf_1 + SaR(1 - f_1)f_2 + (1 - S)R(1 - f_1)(1 - f_2)f_3$
p(EEEE)	$S(1 - a)^3 + S(1 - a)^2aR + S(1 - a)aR(1 - f_1) + SaR(1 - f_1)(1 - f_2)$ $+ (1 - S)R(1 - f_1)(1 - f_2)(1 - f_3)$

Note. Each probability in the left column appears in the empirical likelihood function. In the likelihood function for the trace-integrity model, these probabilities are replaced by the corresponding expression in the right column. C = correct response; E = incorrect response; S = the probability of storage failure; R = the probability of retrieval failure of information in storage; r_1, r_2, and r_3 = the probability of two, three, and four consecutive successes, respectively; f_1, f_2, and f_3 = the probability of a success following: one, two, and three consecutive errors, respectively; a = the probability that information not in storage is redintegrated to a level above zero recall.

(*Appendixes continue*)

both participants and items within each condition of interest), inserting these numbers in the relevant exponents in Equation A2, and maximizing the function using a standard computer optimization routine (e.g., SIMPLEX). The optimal solution yields numerical estimates of the model's nine parameters as well as the value of the likelihood function $\boldsymbol{L}_9$. This latter value (which is more commonly estimated via the log transform $-2\ln\boldsymbol{L}_9$) is used to evaluate the model's goodness of fit (the fourth step) and to examine hypotheses about between- and within-condition differences in the numerical estimates of parameters (the fifth step).

The fourth step involves evaluating the fit of the model to the data. One accomplishes this by maximizing (using the same log transform as above) Equation A1 for the same data as Equation A2, which yields an estimate of the likelihood of the data before the model was imposed (i.e., with all empirical probabilities free to vary, $\boldsymbol{L}_{15}$). Because Equation A1 exhausts all of the information in the data, the value of $\boldsymbol{L}_{15}$ will always be the maximum likelihood for that data set. Because the trace-integrity model does not exhaust this information (having only 9 degrees of freedom, not 15), the estimated likelihood of Equation A2 will tend to be smaller. Goodness of fit is assessed with likelihood ratio tests that determine whether this difference is statistically reliable. This test takes the form

$$\chi^2(6) = (-2\ln\boldsymbol{L}_9) - (-2\ln\boldsymbol{L}_{15}) \qquad \text{(A3)}$$

and evaluates the null hypothesis that the trace-integrity model fits the data.

Finally, the fifth step involves testing hypotheses about the theoretical processes underlying retention performance, as reflected in the numerical estimates of the model's parameters. Because these parameters are identifiable (see Howe, 2000), they can be used in direct tests of hypotheses concerning between-conditions and within-condition differences in the rates of forgetting and reminiscence as well as the storage and retrieval loci of these differences. The statistical process for testing hypotheses is straightforward, involving a series of likelihood-ratio chi-square tests known as an experiment-wise test, condition-wise tests, and parameter-wise tests. The experiment-wise test, like the omnibus *F* test, evaluates the null hypothesis that, on average, the model's parameters do not vary between conditions. The exact test is given by

$$\begin{aligned}\chi^2[k \times (9) - 9] = [(-2\ln\boldsymbol{L}_{9i}) \\ + (-2\ln\boldsymbol{L}_{9i+1}) + \cdots + (-2\ln\boldsymbol{L}_{9k})] - (-2\ln\boldsymbol{L}_{9\text{pooled}}), \quad \text{(A4)}\end{aligned}$$

where the first term represents the summation of the $-2\ln\boldsymbol{L}_9$ values for each of the *i* through *k* individual conditions in the experiment and the last term represents the single $-2\ln\boldsymbol{L}_9$ likelihood value found by pooling the data from all *k* conditions. As there are nine degrees of freedom involved in each of the *k* terms in the first part of the expression, the asymptotic chi-square distribution of the experiment-wise test has $k \times (9) - 9$ degrees of freedom.

The condition-wise test, like the *t* test, evaluates the null hypothesis that, on average, the model's parameters do not vary between specific pairs of conditions. The exact test is given by

$$\chi^2(9) = [(-2\ln\boldsymbol{L}_{9i}) + (-2\ln\boldsymbol{L}_{9j})] - (-2\ln\boldsymbol{L}_{9ij}), \qquad \text{(A5)}$$

where the first term represents the summation of the $-2\ln\boldsymbol{L}_9$ values for conditions *i* and *j* and the final term represents the single $-2\ln\boldsymbol{L}_9$ value found by pooling the data for the *ij* conditions. Because there are always 18 degrees of freedom associated with the first term and 9 degrees of freedom for the last term, the asymptotic chi-square distribution of the condition-wise test has 9 degrees of freedom.

For those pairs of conditions that differ, the parameter-wise test evaluates the null hypothesis that the numerical estimate of a particular parameter does not vary between those two conditions. The exact test is given by

$$\begin{aligned}\chi^2(1) = [(-2\ln\boldsymbol{L}_{9i}) + (-2\ln\boldsymbol{L}_{9j})] \\ - [(-2\ln\boldsymbol{L}'_{9i}) + (-2\ln\boldsymbol{L}'_{9j})], \quad \text{(A6)}\end{aligned}$$

where the first term is the same as in Equation A5 and the second term represents the joint likelihood of the data from the two conditions (i.e., minimizing $\boldsymbol{L}'_{9i}$ and $\boldsymbol{L}'_{9j}$ simultaneously), subject to the single restriction that the parameter being tested (e.g., *S*) assumes the same value in the two conditions. This restriction results in an asymptotic test statistic of $\chi^2(1)$.

Finally, there are two types of within-condition hypotheses that can be evaluated, namely, numerical equivalences (e.g., $a = 0.0$; $S = R = .1$; $r_1 = r_2 = 1.0$) and algebraic relationships (e.g., $S > R$; $r_1 < r_3$). Both types of hypotheses can be tested with the single statistic

$$\chi^2(1) = (-2\ln\boldsymbol{L}_{9i}) - (-2\ln\boldsymbol{L}_{8i}), \qquad \text{(A7)}$$

where the first term is simply the likelihood value associated with condition *i* when all of the parameters are free to vary and the last term represents the same condition with a single restriction imposed. Like the between-conditions parameter-wise test, this chi-square test has one degree of freedom.

Appendix B

Concepts Pairings and the Experimenter- and Self-Generated Similarities and Differences

			Self-generated			
	Experimenter-generated		5-year-olds		7-year-olds	
Concept pairs and category label	Similarities	Differences	Similarities	Differences	Similarities	Differences
Airplane–car (vehicles)	Have wheels	Air–fly, road–drive	Have wheels	Wings, wheels	Have wheels	Flies, drive
Apple–banana (fruit)	Grow on trees	Red, yellow	Eat	Round, long	Eat	Peel, or not
Arm–leg (body parts)	Have joints	Hand, foot	Bend	Fingers, toes	Long	Hand, foot
Ball–doll (toys)	Play with	Bounce, rock	Play with	Bounce (doesn't)	Play with	Bounce (doesn't)
Chair–table (furniture)	Made of wood	Sit, eat off of	Legs	Sit, lay	Legs	Sit, lay
Chicken–owl (birds)	Feathers	Walks, flies	Fly	No fly, flies	Fly	Beaks differ
Dog–cat (animals)	four legs	Barks, meows	Tails	Big, small	Tails (pets)	Barks, meows
Hammer–saw (tools)	Used to build	Nails, cuts	Work	Nail, cuts	Work	Nail, cuts
Knife–fork (eating utensils)	Used for eating	Cutting, pick-up	Eat	Sharp, fingers	Sharp ends	Sharp, prongs
Lobster–fish (seafood)	Ocean	Shell, scales	Tails	Claws, swims	Sea	Claws, fins
Mountain–tree (landscape)	Big (tall)	Wood, rock	High up	Rocks, leaves	Dirt under them	Higher, leaves
Pen–pencil (writing)	Printing	Ink, lead	Long	No eraser, eraser	Sharp	No eraser, eraser
Piano–drum (musical instrument)	Makes sounds	Keys, sticks	Sound	Triangle, circle	Play them	Keys, sticks
Shirt–pants (clothing)	Made of cloth	Top, bottom	Wear them	Top, bottom	Wear them	Top, bottom
Spider–bee (bugs)	Lots of legs	Web, honey	Legs	Big, antennae	Legs	Big, antennae
Sun–star (in the sky)	Bright (shine)	Day, night	Circles	Different shapes	Yellow	Morning, night
Toaster–stove (appliances)	In the kitchen	Counter, floor	Kitchen	Toast, hot	Knobs	Toast, hot
Window–door (house parts)	Open	See, walk	See through	Small, big	Open	Glass, knob

Received May 18, 2005
Revision received September 16, 2005
Accepted October 25, 2005 ■