THE DEVELOPMENT OF LONG-TERM RETENTION IN CHILDREN: DIFFERENTIATING AMNESIA AND HYPERMNESIA

ANDREA J. KELLAND, B.Sc.(Hons.)

ISBN 0-315-55016-3

The Development of Long-Term Retention in Children: Differentiating Amnesia and Hypermnesia

by

© Andrea J. Kelland B.Sc.(Hons.)

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science

Department of Psychology
Memorial University of Newfoundland
September 1989

St. John's Newfoundland

Abstract

Although there is a considerable amount of knowledge about how children acquire information, very little is known about how they retain information in memory. Both acquisition and retention are important in cognition and both must be understood to have a more complete picture of cognitive development. Some of the factors responsible for the absence of research in children's long-term retention, as well as the methodological and analytical refinements necessary for studying children's long-term retention, are discussed. A mathematical model of long-term retention, one that partitions forgetting and relearning into storage and retrieval components, is described and applied to an experiment in which grade 2 and 5 children's retention of 3-item clusters was examined. The clusters varied in semantic relatedness (related or unrelated) and in presentation modality (pictures or words) and retention was examined across 2 sessions over different retention intervals (at 2 and 16 days or 16 and 30 days after acquisition). Both forgetting and relearning were observed at retention with changes in performance being due to alterations in both the availability of information in storage and the retrievability of that information. The most prominent developmental difference was found in forgetting, not relearning, with younger children forgetting more than the older children.

Interestingly, regardless of age, storage failure was greater than retrieval failure. The results of this study were interpreted in the context of the recently developed trace-integrity theory of long-term retention in which both the storage and retrieval aspects of forgetting and relearning are combined into a single unified framework.

Acknowledgements

I would like to thank Dr. Mark L. Howe for his encouragement, cajoling and infinite tolerance during the writing of this thesis. His ability to withstand procrastination is admirable and I have no doubt that without him this thesis would not exist.

I appreciate the advice and comments of my committee members, Dr. Mary Courage, Dr. Julia O'Sullivan, and Dr. Michael Stones, who helped me separate the trees from the forest.

Thank-you to the School of Graduate Studies and the Department of Psychology for financial support as a graduate student.

I am grateful to Geraldine Roe of the Roman Catholic School Board for St. John's, principals Mrs. Yvonne Hepditch of Holy Trinity Primary, Mr. Brian Shortall of Holy Trinity School, Mrs. Dorothy Snow of Ecole St. Gerard, Mrs. Margaret Rumsey of St. Peter's Primary and Mr. William Martin of St. Peter's School for their cooperation. A special thank-you is given to the children who participated in this research for their effort and enthusiasm.

To my parents, who have supported me emotionally and financially throughout, for always believing in me.

Table of Contents

List of Tables

List of Figures

Introduction

In order to have a thorough knowledge of children's cognitive development both the acquisition and long-term retention of information must be understood. While much is known about children's acquisition of information, it is only recently that researchers have focused on the study of long-term retention processes (see Howe & Brainerd, in press). The lack of research on children's retention processes would not be of particular concern if it could be assumed that factors affecting acquisition (either positively or negatively) would have similar effects at retention. Consistent with this supposition, some researchers have reported no differences in the effect of variables (e.g., pictures, words, numbers) at acquisition and retention and, surprisingly, no developmental interactions (i.e., age differences) at retention (Fajnsztejn-Pollack, 1973; Hasher & Thomas, 1973; Lehman, Mikesell, & Doherty, 1985: Sophian & Perlmutter, 1980).

However, recently it has been argued that a number of uncontrolled factors (e.g., level of learning at acquisition, separation of forgetting and relearning) exist in previous research that may have obscured the presence of developmental interactions (see Howe & Brainerd, in press). Although detailed discussion of these factors will be deferred until later, it is important to note that in the few studies in which such factors have been controlled, differential effects of variables have been reported at acquisition and retention and developmental (Age X Retention) interactions have been observed (Brainerd, Kingma, & Howe, 1985; Brainerd & Reyna, 1989; Howe, 1987). For instance, Brainerd et al., (1985) conducted several experiments with second and sixth graders and found age differences at retention and several asymmetries between acquisition and retention. Categorized word pairs were acquired more quickly by the older children but the younger children remembered more after 1 week. Second graders acquired unrelated word pairs faster than categorized pairs, but the retention rate was the same for both lists. Sixth graders learned unrelated pairs faster than categorized, but retained them more poorly. Finally, word typicality had opposite effects at acquisition and retention for both younger and older children. Differences at acquisition between categorized and unrelated word pairs were larger when the items were typical than atypical. But differences at retention were larger when the items were atypical than when they were typical. From these few examples, then, it would seem that there are important developmental changes in children's long-term retention processes that are not predictable from the research on the development of their acquisition processes.

If variables affect acquisition and retention differently, it follows that different theoretical mechanisms may underlie these two memory processes. Clearly, if progress is to be made toward understanding these mechanisms, analytical and methodological problems in existing long-term retention studies need to be corrected. The purpose of this thesis is to use one such corrective procedure (Howe & Brainerd's, in press, trace-integrity model) to examine the development of retention processes in young children (grades 2 and 5). I begin by outlining the paradigm used to investigate retention processes and define the factors that control performance on these tasks. The literature on children's long-term retention is then reviewed and a detailed discussion of methodological and measurement problems associated with this research is provided. Finally, a solution to these problems is presented and used to analyze the long-term retention data obtained in the present research.

Components of Long-Term Retention

The general paradigm for most long-term retention studies involves presentation of material (words, numbers, pictures) to be learned over one or several study trials. After an interval ranging from minutes to weeks after acquisition, subjects receive one or more retention test trials without further study opportunities. Retention performance is usually measured by comparing total recall(recognition) at the end of acquisition with that on the first retention test and, if multiple retention tests are administered, between the first and subsequent retention tests.

Patterns of performance can be described by the use of two global constructs, amnesia and hypermnesia. Amnesia is defined as a <u>net reduction</u> in the number of items recalled(recognized) following the retention interval or across the retention test trials. Hypermnesia is a <u>net increase</u> in the number of items recalled(recognized) following the retention interval or across the retention test trials. Whether amnesia or hypermnesia occurs depends on two other variables, forgetting and reminiscence, that operate at the level of the individual item. Forgetting refers to a failure to recall(recognize) an individual item that was previously recalled(recognized). Reminiscence refers to recall(recognition) of an individual item that was not recalled(recognized) on a previous test. The term <u>relearning</u> will be substituted hereafter for reminiscence to be compatible with the discussion of the model-based findings presented later.

Global performance on long-term retention tests can be one of two types. For amnesia (net reduction) to be present, the amount of forgetting must be greater than the amount of relearning, resulting in fewer items recalled overall. Hypermnesia (net increase) would result from more relearning than forgetting, producing an improvement in net recall (Brainerd & Reyna, 1989; Howe & Brainerd, in press). Both of these results can occur following the retention interval or during the retention test itself if more than one test trial is administered. The only exception to this is that no hypermnesia can be found over the first retention interval (i.e., between the end of acquisition and the first retention test) if criterion learning is used. Obviously, if recall is perfect at the end of acquisition, no improvements in performance will be found on the first retention test.

To illustrate, suppose a set of 20 3-item clusters is learned to criterion and after an initial retention interval of say 2 days, 4 test trials are administered. If 12 clusters are recalled on the first test trial and 10 are recalled on the fourth test trial then amnesia has occurred during the retention test. Alternatively, if 15 clusters are recalled on the fourth test trial then hypermnesia has occurred. However, because in both of these cases concern is focused on global recall, no consideration is given to which particular clusters are recalled. It is only at the level of forgetting and relearning that individual items are of concern; that is, if cluster number 10 is recalled on test trial 1 but not on test trial 4, then it is considered to be forgotten. If cluster number 10 is not recalled on test trial 1 but is recalled on the fourth test trial then it has been relearned. Note that while this example considered only recall on test trials 1 and 4, all 4 trials of the retention test, as well as the what occurs over the retention interval, are considered when assessing amnesia, hypermnesia, forgetting and relearning.

Empirical Issues

As mentioned, most of the research in the area of children's long-term retention has produced little in the way of developmental differences. Any differences that were found tended to be small in absolute magnitude. For example, Fajnsztejn-Pollack (1973) found no age differences between 5-to 16-year-olds in amnesia for pictures over short (2 weeks) or long (48 weeks) retention intervals. Rogoff, Newcombe, and Kagan (1974) also found no age differences in amnesia for 4-, 6-, and 8-year-olds after a retention interval of 1 week. Lehman et al. (1985), after examining the long-term retention of information about presentation modality, concluded that the children they tested did not forget more than the young adults. Finally, Hudson and Nelson (1986) examined the effects of familiarity on children's (3-, 5-, and 7-year-olds) autobiographic memory recall and found that even preschool children remembered events accurately. They suggested that children and adults may store and retrieve autobiographic events in a similar fashion.

While the lack of developmental differences is counterintuitive, it is likely that this is due to uncontrolled factors rather than nonexistent differences. For example, most of these studies used recognition tasks at long-term retention. On average, recognition tasks are less sensitive measures of developmental shifts in children's memory than are recall tasks (Howe & Brainerd, in press). As well, only one or a few study trials were given at acquisition, so that the level of original learning at acquisition was not equated across the different ages studied.

As with the amnesia/forgetting research, the development of hypermnesia/relearning in children has not received much experimental attention. Early in this century, some experimenters reported an increase in memory with repeated recall attempts. For example, Ballard (1913) found that children's recall of prose improved across repeated test trials even though no additional study opportunities were administered following acquisition. Interestingly, Ballard (1913) found that this result was inversely related to age, such that <u>younger</u> children displayed more hypermnesia than older ones. Several other researchers have reported similar trends (see Piaget & Inhelder, 1973, for a review).

Vertes (1931/32) tested 6- to 18-year-olds for retention of word pairs immediately, 1 day, and again 1 week after acquisition. The older children retained more than the younger ones on the immediate test, forgot less at 1 day, and improved their recall (relative to the immediate recall test) at the 1 week test. Unlike previous researchers, Vertes found that recall improvement at 1 week was restricted to the older children (10 years and up) while younger children displayed amnesia at 1 week.

This pattern of conflicting results, along with the lack of clear developmental trends for either amnesia or hypermnesia, may have contributed to the decline in research, from the 1930s until recently, on children's retention processes. Methodological problems also plagued this early research and subsequent attempts at replication following correction of these problems proved futile (see Howe & Brainerd, in press). In addition to such problems as the type of task used (recognition) and incomplete learning at acquisition, early studies in which hypermnesia had been reported used within-subjects designs, which confound retention interval with prior testing. That is, if various retention intervals are being studied for evidence of hypermnesia, subjects tested at longer intervals are also tested at the shorter one(s). To illustrate, if the long-term retention intervals are 1 and 2 weeks after acquisition, subjects tested at 1 week are also tested again at 2 weeks. Between-subjects designs permit separation of the retention interval and the retention test - some subjects are tested after a short interval while others after a longer one, but neither has received a prior test and, hence, no prior practice. Again using the 1 and 2 week example, some subjects are tested for retention at 1 week after acquisition, while other subjects are tested at 2 weeks. When between-subjects designs such as these were conducted, no increases in recall were observed (see Payne, 1987). For example, Ammons and Irion (1954) found that whereas a within-subjects manipulation produced increased recall, the between-subjects manipulation produced lower net recall with longer retention intervals.

A further dilemma centers around the source of hypermnesia. Is it due to improvements in relearning with age or due to age reductions in forgetting? Paris (1978) presented grade 2 and 6 children with a list of categorically related words and then gave them three free recall tests in succession. Both grades recalled new words on each successive recall trial (relearning), however the grade 2 children forgot more words previously recalled as well. The resulting developmental interaction in performance was due to a decline in forgetting with age, rather than an increase in relearning.

The hypermnesia/relearning component of children's long-term retention is beginning to become a subject of research again, after many years of disinterest. Shaw (1985, cited in Richardson, 1985) found greater hypermnesia for concrete than abstract material, although no developmental interaction was obtained. More recently, however, Brainerd and Reyna (1989) found a developmental interaction such that while amnesia and forgetting decreased with age, hypermnesia and relearning increased between grades 2 and 6.

It would appear, then, that the area of long-term retention in children's cognitive development needs a great deal of work. Methodological and analytical improvements are necessary to correct the problems of previous research and to uncover any developmental interactions that may exist.

Conceptual Issues

To uncover the presence and direction of developmental interactions in amnesia and hypermnesia, the variables that affect forgetting and relearning must be isolated. As mentioned, methodological problems associated with previous research may have obscured the existence of any interaction. Specifically, these problems include failures to: (a) equate level of learning at acquisition, (b) separate forgetting from relearning, and (c) isolate changes in storage and retrieval processes that contribute to long-term retention performance.

To begin, consider the problem of equating the level of original learning. In most long-term retention experiments, there is usually only one, or at most a few, study trials administered at acquisition. Because of individual differences in item learnability, recall on long-term retention tests is confounded with the level of original learning. Worse, in developmental studies where fixed-trials designs are used, level of learning and age are confounded because older children tend to learn any list faster than younger children (e.g., Howe, Brainerd, & Kingma, 1985). Because learning curves are negatively accelerated, these discrepancies will be greater the fewer the number of study trials. Clearly, failures to equate level of learning leave open the possibility that observed levels of amnesia and hypermnesia simply reflect differences in the level of initial learning rather than differences in item forgetting or relearning. Further, the ambiguity noted earlier concerning the existence and direction of hypermnesia may be the result of variation in the numbers of study trials used across the different experiments.

Developmentally, if (a) forgetting increases and relearning decreases as level of learning at acquisition deceases, and (b) forgetting decreases and relearning increases with age, then Age X Retention interactions may not be detected despite their existence when level of learning is not controlled. Again, because learning curves are negatively accelerated, learning tends to become equated across age and lists as the number of study trials increases. The most effective solution, therefore, is to require subjects to meet a stringent acquisition criterion of 2 or 3 errorless passes through the list. Any residual differences in learning at the end of acquisition can be adjusted by fitting Markov models to the acquisition data (e.g., Howe & Hunter, 1986) and "correcting" subjects' long-term retention scores (also see Howe & Brainerd, in press). When a strict criterion of 2 or 3 errorless trials is required at acquisition, this correction is usually very small and only minimal adjustments to the retention data are necessary.

The second problem concerns the separation of forgetting and relearning components of long-term retention. As mentioned, amnesia and hypermnesia can be decomposed into forgetting and relearning components where amnesia signals greater forgetting than relearning, and hypermnesia indicates greater relearning than forgetting. Because amnesia and hypermnesia are aggregate variables, they refer simply to global outcomes (net decrease or increase in total recall) on long-term retention tests. As we seek to differentiate the underlying processes that make up amnesia and hypermnesia, it is paramount that we obtain independent estimates of the contributions of forgetting and relearning to total performance scores. In the current investigation, forgetting and relearning will be examined by analyzing the recall of individual items.

This leads directly to the third problem, namely, whether long-term retention performance is the result of changes in the availability (what is stored) and/or the accessibility (what is retrievable) of the memory trace. If a strict criterion is used at acquisition then it can be safely assumed that the material has been stored in memory and is highly retrievable when retention is tested immediately (e.g., Brainerd, Howe, & Kingma, 1982; Howe & Brainerd, in press). On long-term retention tests, forgetting and relearning may be due to changes in what is stored, how it is retrieved, or both. Because there is considerable theoretical controversy concerning the role of storage and retrieval processes at retention these alternatives are discussed in greater detail below.

Storage and Retrieval Interpretations of Retention Performance

The usual explanation of forgetting is that it is due to retrieval failure. Ceci, Ross, and Toglia (1987) stated that in general the current belief in cognitive development is that memories are enduring and that once a memory trace is formed it becomes permanent (see also Loftus & Loftus, 1980). Retrieval techniques such as free association, hypnosis, even Penfield's (1969) brain stimulation experiments, all of which may produce memories of seemingly forgotten information, are given as support for the permanence of memory. Loftus and Loftus (1980) reviewed examples in the cognitive literature of the memories produced by such techniques and found that many are actually reconstructions rather than retrieval of intact memories. They suggested that memory traces may be labile rather than permanent and, therefore, susceptible to loss or alteration. This and other explanations of forgetting have recently been postulated and debate has begun over whether information is actually lost from memory (no longer stored), is in memory but just not retrievable at the time, or is altered so that the original trace varies with respect to its original integrity in storage. To properly explain what the processes known as forgetting and relearning are, the issues of inaccessibility, irretrievability and trace alteration must be addressed.

In terms of forgetting, research in eyewitness testimony and the leading-questions paradigm has led to considerable debate over the reason(s) for the 'forgetting' of the original information. Briefly, the leading-questions paradigm involves presenting (visually and/or auditorily) subjects with an event and sometime later providing misleading information about the original event (e.g., asking "what colour was the Stop sign?" when in fact it was a Yield sign in the original presentation). Some researchers have found that the misleading information affects subjects' recall of the original event such that the new information is provided as the original. The work of Loftus and her colleagues (e.g., Loftus, 1979b; Loftus, Hoffman, & Wagenaar, in press; Loftus and Loftus, 1980) has lead to the contention that the memory trace can be distorted or changed such that the original trace is no longer retrievable. McCloskey and Zaragoza (1985), on the other hand, suggest that the original memory trace can coexist with a changed trace and that either is potentially retrievable. Ceci et al. (1987) seem to prefer a somewhat middle ground, where memories may be enduring but it may also be possible to transform them, making the original traces inaccessible. This latter explanation is given for their finding that younger (3- and 4-year-old) children were more susceptible to biased information than older children and adults. They suggested that younger children may be more susceptible because they forget more of the original information leaving less on which to base subsequent recollections. However, this account may not be tenable because they were unable to detect differences in forgetting across the age groups.

Currently, the interpretation of this class of findings is very contentious, with hypotheses about what happens when misleading information is introduced ranging from memory impairment to the coexistence of the original and misleading information (see Belli, 1989; Tversky & Tuchin, 1989; Zaragoza & McCloskey, 1989). However, one thing is clear, memory traces are not immutable and that perhaps mutability interacts developmentally (for a review, see Loftus et al., in press). In fact, changes in trace mutability may be related to developmental differences in trace strength (Howe & Brainerd, in press). Consider Brainerd and Reyna's (1988) explanation of Ceci et al.'s (1987) results. They point out that recognition tasks tend to be insensitive to measuring the development of forgetting and therefore the possibility of forgetting affecting suggestibility cannot be dismissed. Because age invariance in forgetting has recently been dispelled (Brainerd et al., 1985; Howe, 1987; Howe & Brainerd, in press), the finding that young children forget more of the original information than older children does not necessarily lend support to the altered trace hypothesis postulated by Ceci et al.(1987). Rather, as Brainerd and Reyna (1988) point out, age-related factors (e.g., rate of learning at acquisition) that influence trace strength during storage and retrieval can lead to the greater suggestibility of the younger children. If trace strength is viewed as a continuum, where the orignal intact trace is at one end and a completely altered trace is at the other, there are innumerable possibilities for changes in the trace that may be related to factors such as age. Ceci, Toglia, and Ross (1988) agreed with Brainerd and Reyna's (1988) critique but extended the argument to support their trace alteration hypothesis. Specifically, they posited that the weakening of the original trace may exacerbate its alteration by misleading questions. Whatever the outcome of the controversy over what happens to memory traces, it is clear that a pure retrieval explanation for long-term retention findings is untenable.

Like forgetting, the locus of hypermnesia and relearning is unclear. Piaget and Inhelder (1973) hypothesized that alteration of the original memory trace was responsible for improvements in recall even after a 6 to 12 month retention interval. They refer to their findings of improved performance across a variety of cognitive tasks as due to the reconstruction or transformation of the original information. The idea that improvements in recollection can be due to changes in the actual contents of memory traces (storage), and not just improvements in the retrieval conditions at retention, is consistent with modern theories concerning the operation of working memory (e.g., Brainerd & Reyna, in press).

While it appears to be difficult to separate the effects of storage and retrieval at long-term retention, both in terms of forgetting and relearning, the debate over which is responsible makes it clear that both storage and retrieval processes must be considered in studies of long-term retention. Trace-integrity theory provides explicit mechanisms to deal with these problems (Howe & Brainerd, in press). In this theory, Howe and Brainerd (in press) suggest that while storage- and retrieval-based forgetting may be different memory processes, they can also be viewed as two components of a single process that lies on a continuum. If the original trace consists of a set of well-encoded features, then the integrity of the feature set, and thus the trace, should be the primary determinant of how accurately the trace is recalled. Disintegration of the bonds holding the feature set together is related to both storage- and retrieval-based forgetting. The beginning of trace disintegration is associated with retrieval-based forgetting (trace inaccessibility) with further disintegration resulting in storage-based forgetting (trace unavailability). In other words, when the trace is just starting to erode in memory, failure to recall is related to a problem of retrieval. As trace erosion progresses, recall failure is a result of the trace being unavailable from storage. Of course, storage failure does not mean the trace is necessarily lost from memory, but rather, its level of integrity is such that recall occurs with probability zero.

As forgetting is associated with disintegration and decrements in recall, relearning is associated with redintegration (Horowitz & Prytulak, 1969) of the trace and improvements in recall (i.e., hypermnesia, reminiscence). Redintegration refers to a 'rebonding' of the features of a trace that has gradually disintegrated so that it becomes a coherent unit again. Increased recall across test trials is most often explained by improvements in retrievability due to practice effects (Runquist, 1986a, 1986b, 1987). However, as with forgetting, if storage and retrieval are viewed as elements of the same phenomenon, storage relearning, or restorage, should be considered along with retrieval relearning. Restorage refers to redintegration of traces that have fallen to the zero recall threshold. That is, featural activation spreads throughout the trace, reactivating the unit in memory, and permitting recall to cross the zero threshold. Similarly, retrieval relearning consists of featural reactivation and spread until appropriate retrieval mechanisms are reinstated and the trace becomes accessible for recall. Retrieval relearning is possible after successful recall or after an error in recall on previous test trials.

To summarize, storage and retrieval can be thought of as two components of a single process and both must be considered when examining forgetting and relearning. Further, the disintegration/redintegration hypothesis of Howe and Brainerd (in press) postulates that recall on retention tests is determined by the strength, or amount of featural integration, of the trace. Forgetting is due to the weakening of the trace, or featural disintegration, with early disintegration related to retrieval-based forgetting and further disintegration resulting in storage-based forgetting. Relearning (both restorage and retrieval relearning) is a result of a reactivation of some features of the trace, with the spread of featural activation continuing until the trace is redintegrated and recalled.

Model-Based Analysis of Long-Term Retention Performance

The mathematical model associated with the trace-integrity theory (Howe & Brainerd, in press) will be used to factor the contributions of restorage, retrieval relearning, storage failure, and retrieval failure in the present research. Definitions of the model's parameters are provided in Table 1. This trace-integrity model is designed to give independent estimates of the availability (in storage) and the accessibility (retrievability) of the memory trace after the retention interval, and of relearning (restorage and retrieval-based) during the retention test itself. These independent estimates are obtained by separating forgetting and relearning through the use of a stochastic model defined over an outcome space that consists of 16 unique combinations of correct (C) and incorrect (E) responses across the four test trials of each retention test. The relevant equations are provided in Table 2.

The nine independent parameters of the long-term retention model (see Table 1) are divided into two that measure forgetting and seven that measure relearning. The forgetting parameters are $\underline{S}$, for storage-based forgetting, and $\underline{R}$, for retrieval-based forgetting. $\underline{S}$ gives the unconditional probability that an item is unavailable following the retention interval and $\underline{R}$ gives the conditional probability that an item that is in storage is not accessible.

In terms of relearning, there is a single restorage parameter, $\underline{a}$, which measures the conditional probability that information that was unavailable after the retention interval is restored (through processes that redintegrate the trace) during test trials. The remaining six parameters all measure retrieval relearning, three of which assess relearning following successful recall and the other three measure relearning following an error. The success relearning parameters, $\underline{r}_1$, $\underline{r}_2$, and $\underline{r}_3$, measure the probability of successful recall following one, two, or three preceding successes, respectively. The error relearning parameters, $\underline{f}_1$, $\underline{f}_2$, and $\underline{f}_3$, measure the probability of successful recall following one, two, or three consecutive errors, respectively. Comparing the values of the $\underline{r}$'s to the $\underline{f}$'s gives an indication of when more relearning occurs, after a success or after an error, and consequently which is more important in re-establishing trace retrievability (see Howe & Brainerd, in press).

To summarize, this model uses a mathematical procedure for separating the forgetting and relearning components of both amnesia and hypermnesia and determines whether the source of these contributions are at storage and/or retrieval. The forgetting parameters ($\underline{S}$ and $\underline{R}$), in combination with the relearning parameters ($\underline{a}$, $\underline{r}$'s and $\underline{f}$'s) will permit the partitioning of the origin(s) of any developmental variation in amnesia or hypermnesia. With these estimates in hand, a more complete discussion about whether net decrements and/or improvements in recall are due to changes in trace accessibility, trace availability, or both, can ensue.

Present Research

The dearth of results with children in all areas of long-term retention provided the impetus for the present research. A developmental comparison of retention with repeated testing over varying intervals should provide some insight into the variables affecting forgetting and relearning. This experiment involves children learning material to a strict criterion of three consecutive errorless test trials. Three manipulations were used to better understand children's long-term retention. First, a general analysis of previous research in hypermnesia (at least with adults) might lead to the conclusion that pictures produce more hypermnesia than words (e.g., Erdelyi & Becker, 1974; Roediger & Payne, 1982) and that these modality manipulations produce greater hypermnesia than semantic ones (Belmore, 1981). However this conclusion is premature because a direct comparison between modality and semantic factors has rarely been made within the same experiment. In the present study such comparisons will be made by having different groups of subjects learn clusters of unrelated pictures, unrelated words, related pictures or related words. In this way the relative magnitude of the effect of these factors on amnesia and hypermnesia can be directly evaluated.

A second series of comparisons was included to determine what effects repeated testing and time of test would have on retention performance. In order to separate the effect of length of retention interval from repeated testing, a between-subjects manipulation is necessary. This was achieved by using two retention intervals with the time of the first test being varied. For this experiment, half of the subjects were tested at 2 days after acquisition and again at 16 days, while the other half were initially tested at 16 days (denoted as 16′ to avoid confusion with the 16-day second test of the other group) and then again at 30 days. The three effects that can be evaluated from these manipulations are, (a) differences in retention performance as a function of the length of the initial interval (2- v 16′-days), (b) the effect of the presence versus the absence of a previous test on retention performance at 16 days (16- v 16′-days), and (c) the effect of the time of first test (early 2-days and late 16′-days) on the second retention test (16- and 30-days).

Finally, a developmental comparison was included. Although it is well known that developmental differences exist when modality and semantic relatedness are manipulated at acquisition, it is not clear that the same effects are found at retention. In order to remedy this situation, and in order to uncover developmental differences in young children's amnesia, forgetting, hypermnesia, and relearning, elementary school children in both grades 2 and 5 participated in the experiment.

Method

Subjects.

One hundred and sixty grade 2 (Mean age = 7 years, 5 months, SD = 4 months) and 160 grade 5 students (Mean age = 10 years, 4 months, SD = 4 months) were tested. An equal number of males and females participated at each grade level and parental consent was obtained for each child's participation.

Materials.

Subjects learned a set of 6 three-item (picture or word) associative clusters, each cluster being presented on a separate index card (see Appendix A). All items for the clusters were concrete concepts obtained from the Snodgrass and Vanderwart (1980) norms and, with the addition of the Battig and Montague (1969) and Toglia and Battig (1978) norms, were matched on concreteness, familiarity, typicality, and picturability. There was a total of four lists, two related and two unrelated. The related lists consisted of 3-item clusters in which each of the three items were obtained from the same category. For half the subjects, the related clusters were presented in pictorial format and for the other half, they were presented as words. The unrelated lists consisted of three items for each cluster being obtained from different categories, where again one of lists was presented as pictures and the other one as words. The first item of each cluster was designated as the cue, the other two members of the cluster being the targets.

Procedure.

Subjects were randomly assigned to the different conditions with the caveat that there be an equal number of males and females in each group. Eighty subjects in each grade were given word clusters and the other eighty were given pictures. Within each cluster group, half of the subjects learned related clusters and half unrelated clusters. Subjects were further divided into two different retention interval groups. Twenty in each list condition were tested at 2 days and again at 16 days after acquisition, while the other twenty were tested at 16 days and again at 30 days after acquisition.

Subjects were tested individually using a standard study-test procedure. A study trial was given followed by two test trials in succession. Thereafter the order was study trial - test trial until the subject learned all six clusters to a criterion of three consecutive errorless trials, with a maximum of 25 acquisition trials allowed. Each cluster was presented separately at a seven second rate while being read aloud by the experimenter. The verbalization was included in consideration of the reading ability of the subjects, especially the grade 2's, and to make sure that no differences occurred with the labelling of the pictures.

Clusters were randomly presented to avoid serial position effects. In order to avoid short-term memory effects the last few items on a study or test trial were never among the first few items on the next study or test trial. On test trials, the cue was presented and the subject was to respond with both targets. Guessing was encouraged and subjects were told to respond even if they only remembered one of the two targets.

The long-term retention tests consisted of four test trials with no further opportunity for study of the entire cluster. The same controls used at acquisition to prevent short-term memory and serial position effects were used at retention. At both acquisition and retention, responses were recorded individually so that if only one target was retrieved it was noted, although for the purposes of scoring, a correct response consisted of recall of both targets. Later examination of the recall of the individual targets showed very little partial recall; subjects either recalled both of the targets or neither. Consistent with previous cluster research (Howe, 1985), analysis of partial recall produced the same pattern of results as analysis of the entire cluster. Because interest is focused on recall of the entire memory trace, results of the analyses of the entire cluster will be reported.

Results

Initially an analysis of variance (ANOVA) was conducted to obtain global findings for amnesia and hypermnesia. These results are reported first, followed by the findings obtained by applying the trace-integrity model to the data. All results were significant at $p < .01$.

Summary analyses

Analyses were conducted on recall of the entire cluster obtained during the long-term retention test trials. The mean number of correctly recalled clusters for each of the long-term retention (LTR) sessions is given in Table 3. These data were analyzed using a 2 (grade: 2 v 5) x 2 (LTR session: 1 v 2) x 2 (semantic: unrelated v related) x 2 (modality: pictures v words) x 2 (retention interval: 2-16 v 16'-30) x 4 (trials) analysis of variance (ANOVA).

Significant main effects were found for the semantic ($F(1,304) = 356.39$), retention interval ($F(1,304) = 62.52$) and trials ($F(3,912) = 37.95$) factors. As expected, related clusters were retained better than unrelated clusters (Mean = 5.41 v 3.18) and the early retention interval (2-16 day) produced better recall than the later (16'-30 day) interval (Mean = 4.76 v 3.83). Finally, post hoc Newman-Keuls tests on trials showed Trial 1 performance was poorer than Trials 3 and 4, and Trial 2 performance was inferior to Trial 4.

A two-way interaction was found for retention session x retention interval ($F(1,304) = 37.50$). Post-hoc tests showed that while recall for the 2-16 day interval was greater than 16'-30 day, there was no difference between sessions 1 and 2 for 2-16 day but there were significant differences between 16'- and 30-day tests (see Figure 1). A three-way interaction for retention session x retention interval x trials ($F(3,912) = 3.96$) was the only other higher-order effect found (see Figure 2). Three important results were revealed by post-hoc tests. First, hypermnesia was found across test trials, with improvements in recall found particularly for the 2-16 day retention interval. Second, performance on the 2-16 day interval was greater than for the 16'-30 day interval. Finally, with trials included as a factor, recall declined between 2 (Trial 4) and 16

(Trial 1) days but remained stable between the last 16'-day trial and the first 30-day trial.

From these analyses it would seem that retention performance was affected by the time of testing and semantic relatedness. Retention performance improved across test trials (hypermnesia) within both retention sessions (2-16 days and 16'-30 days). No modality or age differences were found. Below, the loci of these effects (storage, retrieval, forgetting, relearning) will be determined using the trace-integrity model.

Model-based analyses

Before using the long-term retention model, it has to be determined that the model provides an adequate account of the data. Goodness of fit (see Appendix B) was evaluated using standard likelihood-ratio procedures (see Howe & Brainerd, in press, Eqs. 1-2). None of the 32 goodness-of-fit statistics calculated for the present data resulted in the rejection of the null hypothesis, a finding that indicates that the model adequately captured the data (see Table 4). Because the nine parameter long-term retention model fits the data, its parameters can be used to investigate hypotheses concerning the locus of amnesia and hypermnesia. The parameter estimates for the model are given in Table 5.

Hypothesis-testing was conducted in three phases. First, an experimentwise test was performed to evaluate the null hypothesis that the parameter values were not different between conditions in the experiment as a whole. This test is analogous to an omnibus F test and the result [$X^2(279) = 1429.41$] indicated that differences did occur. Second, a series of conditionwise tests, analogous to t-tests, were conducted to determine which pairs of conditions differed. A total of 88 $X^2(9)$ conditionwise tests were conducted (see Table 6): 16 conditionwise tests were conducted to evaluate each of the developmental, semantic, modality, and retention session (1 v 2) effects, and eight tests were conducted to evaluate each of the effects of a preceding v no preceding retention test on retention at 16 days (16 v 16′), the effect of timing of the initial retention test (2- v 16′- days), and of the timing of the second retention test (16- v 30-days). Third, parameterwise tests were used to evaluate the null hypothesis that the value of a specific parameter did not differ between pairs of conditions that differed significantly in Table 6. For each significant conditionwise test, each of the nine parameters of the model was compared. Because 67 of the 88 conditionwise tests were significant, 603 (67 x 9) parameterwise tests were conducted. Due to the very large number of parameterwise tests that were conducted, only those that were significant are reported. Rather than simply listing the considerable number of $X^2(1)$ differences for the parameterwise tests, it is customary to summarize the significant findings according to the relevant effects being studied, namely developmental, modality, semantic, time of test, and test-retest effects. Again, all of the effects summarized below were significant at $p < .01$.

Developmental Effects

Forgetting. Regardless of whether one examines forgetting or relearning, the most predominant effect overall was greater storage-based forgetting. Specifically, grade 2's exhibited more storage-based forgetting than grade 5's, who exhibited more retrieval-based forgetting. For storage-based forgetting, all of the comparisons indicated greater forgetting for the grade 2 than grade 5 children. Three of the four unrelated pictures comparisons (16-, 16'-, and 30-day tests) and one of the unrelated words (30-day test) were significant. No forgetting differences were found for the related lists. Grade 5 children exhibited more retrieval-based forgetting than grade 2 children with unrelated pictures (16- and 16'-day tests) and unrelated words (30-day). Only one retrieval-based forgetting comparison indicated higher failures of this sort for the grade 2's (unrelated words, 16'-day test).

Relearning. Very little relearning, either restorage or retrieval relearning, was found for either grade. Significantly more restorage was shown by the grade 2's for unrelated pictures, 2-day test, and grade 5's for related words, 16-day test. Further evaluation of restorage can be obtained by examining the probability of its occurrence on any test trial ($\underline{a}$) and the cumulative probability that it occurred on one of the four test trials [$\underline{a} + \underline{a}(1-\underline{a}) + \underline{a}(1-\underline{a})^2 + \underline{a}(1-\underline{a})^3$]. The average restorage rates for the younger (.05) and older (.04) children, and the cumulative restorage rates (.19 and .16, respectively) indicate a lack of overall developmental difference at restorage.

Success-contingent retrieval relearning can be evaluated by comparing the values of the $\underline{r}$'s to the initial probability of item retrieval ($1-\underline{R}$). If the $\underline{r}$'s are larger than $1-\underline{R}$ then success-contingent retrieval relearning has occurred. No developmental differences were found for success-contingent retrieval relearning, either at the level of each condition or in the averaged rates (grade 2, $1-\underline{R}$=.92, $\underline{r}_1$=.94, $\underline{r}_2$=.97, and $\underline{r}_3$=.99; grade 5, $1-\underline{R}$=.88, $\underline{r}_1$=.93, $\underline{r}_2$=.98, and $\underline{r}_3$=.99). For error-contingent retrieval relearning, no consistent developmental pattern was observed, with grade 2's being significantly better for some comparisons ($\underline{f}_1$: unrelated pictures, 16- and 16'-day tests, unrelated words, 16- and 30-day tests, related words, 16-day test;

$\underline{f}_2$: unrelated words, 16-day, related words, 16- and 16′-day; and $\underline{f}_3$: unrelated words, 30-day test), and the grade 5's being better for others ($\underline{f}_1$: unrelated words, 16′-day, related words, 16′-day; and $\underline{f}_2$: unrelated pictures, 2- and 30-day tests). Error-contingent retrieval relearning can also be evaluated by comparing the values of the $\underline{f}$'s to the initial probability of item irretrievability ($\underline{R}$). If the $\underline{f}$'s are larger than $\underline{R}$ then error-contingent retrieval relearning has occurred. Again, as with the success-contingent retrieval relearning, no developmental differences were found for the averaged rates of error-contingent retrieval relearning (grade 2, $\underline{R}$=.08, $\underline{f}_1$=.66, $\underline{f}_2$=.32, and $\underline{f}_3$=.29; grade 5, $\underline{R}$=.12, $\underline{f}_1$=.56, $\underline{f}_2$=.26, and $\underline{f}_3$=.15), although it did tend to decline across trials.

Developmentally, then, forgetting, not relearning, would seem to be the predominant factor differentiating elementary school children at long-term retention. Considerably more storage-based forgetting was exhibited by the younger children while most retrieval-based forgetting occurred with the older children. The average difference for both types of forgetting was greater for the pictures than words and for the unrelated than related clusters. A final important point is that, regardless of age, storage failure was more prominent than retrieval failure.

Modality Effects

Forgetting. Few forgetting differences were found based on the modality of presentation. Only five significant differences were found overall, three for storage-based forgetting and two for retrieval-based forgetting. The only storage-based forgetting differences were confined to more forgetting for the word than picture lists (grade 2, related, 30-day; grade 5, unrelated, 16'-day and related, 30-day). Retrieval-based forgetting differences were also greater for words than pictures (grade 2, unrelated 16-day and related 16'-day).

Relearning. As with forgetting, only a small number of relearning comparisons were significantly different between pictures and words. Those that did occur resulted from the greater relearning of pictures. That is, pictures were restored significantly more than words for both age groups. The trend for grade 2's was pictures generally being restored better than words (average restorage rates .09 v .02 and cumulative .31 v .06, respectively) while no differences were observed between pictures and words for the grade 5's (average restorage .04 v .05 and cumulative .14 v .18, respectively). Grade 2's restored unrelated pictures more than unrelated words on the 2- and 16-day tests and related pictures more than related words on 16'-day test and the only grade 5 difference was greater restorage for related pictures than words on the 16'-day test.

No differences in success-contingent retrieval relearning were observed for either the grade 2's or 5's for the picture/word manipulation. Error-contingent retrieval relearning favoured pictures over words in all significant comparisons for the grade 2's ($\underline{f}_1$: related 16'-day; $\underline{f}_2$: related 16'- and 30-day tests) except one (unrelated words > pictures, $\underline{f}_1$ and $\underline{f}_2$: 2-day test). All the grade 5 differences favoured pictures ($\underline{f}_1$: unrelated 30-day; $\underline{f}_2$: unrelated 30-day, related 16'- and 30-day tests).

In summary, then, modality effects were fairly minimal with those that did occur mostly favouring pictures over words. That is, pictures were remembered better (greater forgetting of words) and were more likely to be relearned if forgotten.

Semantic Effects

Forgetting. The semantic manipulation affected forgetting to a considerable extent, with unrelated material being forgotten more than related in all cases. All of the storage-based and most of the retrieval-based forgetting comparisons were significant. Storage-based forgetting was significantly greater for the semantically unrelated than related lists in all comparisons for both grades. The average rate of storage-based forgetting was exactly the same regardless of whether semantic comparisons involved pictures or words for the grade 2's (unrelated items .47 v related items .08) and only slightly different (not significantly) for the grade 5's (unrelated - pictures, .30, words, .39; related - pictures, .06, words, .10). Retrieval-based forgetting occurred significantly more for unrelated than related pictures and words on the 2-day test for the younger children. For the older children, retrieval-based forgetting was greater for all of the unrelated picture conditions compared to the related pictures, and for the unrelated 16- and 30-day word tests. Thus, storage-based forgetting occurred more frequently than retrieval-based when materials were not semantically related, with the average size of the effect being somewhat greater for the grade 2's (.28) than the grade 5's (.21).

Relearning. Success- and error-contingent retrieval relearning both produced significant comparisons, in some cases being greater for related clusters while in others being greater for unrelated. The most interesting finding was the success-contingent retrieval relearning that occurred over the four test trials, producing hypermnesia.

The only significantly different restorage comparison was for grade 2 unrelated v related pictures, 2-day test. The lack of restorage differences may be due to the high level of retention of the related lists as evidenced by the very low forgetting rates. For the grade 2's unrelated clusters were generally restored better than related (average .08 _v_ .03; cumulative .26 _v_ .12). For the grade 5's, there were virtually no differences between unrelated and related clusters (average .05 _v_ .04; cumulative .18 _v_ .14).

As with the developmental and modality effects, no differences were found for success-contingent retrieval relearning for the grade 2's. However the grade 5's relearned related pictures ($\underline{r}_1$: 16'-day; $\underline{r}_2$: 30-day) and related words ($\underline{r}_1$: 30-day) better than their unrelated counterparts after a previous success. As well, retrieval relearning occurred over the four trials (T) for the unrelated clusters (T_1=.80, T_2=.88, T_3=.96 and T_4=.98).

Error-contingent retrieval relearning differences were found for both grade 2's and 5's. For grade 2's, unrelated words were relearned significantly more than related words after one error ($\underline{f}_1$) at 30-days and unrelated pictures were relearned significantly more than related after two consecutive errors ($\underline{f}_2$) at 16-days. For grade 5's, unrelated pictures were relearned significantly more than related after two consecutive errors ($\underline{f}_2$) at 2-days. Related clusters were relearned significantly better than unrelated for grade 2's on pictures at 2-days ($\underline{f}_1$), 16'-days ($\underline{f}_1$ and $\underline{f}_2$), 30-days

($\underline{f}_2$) and on words at 16'-days ($\underline{f}_2$). For the grade 5's, related words were relearned more than unrelated at days 16 and 30 ($\underline{f}_1$) and related pictures more than unrelated pictures at days 16' and 30 ($\underline{f}_2$). As would be expected, as the number of previous consecutive errors increased, the probability of a success decreased, or conversely, the probability of another error increased (grade 2, $\underline{R}$=.08, 1-$\underline{f}_1$=.34, 1-$\underline{f}_2$=.68, and 1-$\underline{f}_3$=.71; grade 5, $\underline{R}$=.12, 1-$\underline{f}_1$=.44, 1-$\underline{f}_2$=.74 and 1-$\underline{f}_3$=.85).

In summary, forgetting, especially storage-based forgetting, was particularly affected by the semantic manipulation. The relearning that occurred for the unrelated clusters due to success-contingent retrieval relearning produced the sought after hypermnesia - increased net recall over test trials. Interpretation of this must be tempered by the finding of the high level of retention of the related lists as evidenced by the relatively low average rate of forgetting.

Three test and time comparisons were conducted to assess the effects of retesting, the timing of the retention tests, and the effect of a preceding test with time held constant. These are discussed separately, below.

Test-Retest Effects

Comparisons were made within each of the retention intervals (2-16 and 16'-30) to assess the effects of a prior test on a subsequent test.

Forgetting. More forgetting occurred at the second interval (16'-30 days), most of it was exhibited by the grade 2's, with day 16 (both 16 and 16') being the time at which forgetting peaked. Here, grade 2's exhibited greater storage-based forgetting on day 16 than day 2 of the first retention interval for both unrelated pictures and words. No storage-based differences were found in the 2-16 interval for the grade 2's on the related lists or for the grade 5's on any of the unrelated or related lists. In the second interval (16'-30), the only grade 2 difference was greater 16'-day than 30-day storage-based forgetting for unrelated pictures. The only grade 5 difference in the second interval was more storage-based forgetting on 16'-days than on 30-days for unrelated words.

The only retrieval-based forgetting difference in the 2-16 day interval for both grades was found on day 2 for the younger children for unrelated pictures. More retrieval-based forgetting occurred during the later than early retention test. Greater forgetting was exhibited on the first test (16') for unrelated and related words, and on the second test (30) for unrelated pictures.

To summarize, grade 5's showed very few differences in forgetting within either of the retention intervals, the second interval (16'-30 days) produced more forgetting than the first (2-16 days), the greatest amount of forgetting occurred 16 days after acquisition (both 16- and 16'-days), and again, unrelated material was affected the most.

Relearning. The biggest effect for relearning was due to error-contingent retrieval relearning, which occurred more during the later retention interval (16'-30) than the early one (2-16). While grade 2's exhibited error-contingent retrieval relearning at both the early and later intervals, the grade 5's exhibited more during the later interval than at they did at the early one.

Only 2 significant differences were found for the restorage parameter, one for each grade. Grade 2's restored more at 2 days than at 16 days for unrelated pictures. Grade 5's restored more on the second than the first test of the first interval (16-days) for related words. Although few parameter differences were found, trends of cumulative restorage rates were higher restorage on the first (2-16 days) interval (.24 and .23 for grades 2 and 5, respectively) than on the second (16'-30 days) interval (.14 and .08).

Success-contingent retrieval relearning differences were minimal. Grade 2's showed more relearning on day 16' than day 30 ($\underline{r}_1$: unrelated words) and grade 5's showed more relearning on day 30 than on day 16' ($\underline{r}_1$: unrelated pictures). Considerably more error-contingent retrieval relearning occurred. For the early retention interval (2-16), grade 2's relearned more at 16-days ($\underline{f}_1$: unrelated pictures; $\underline{f}_2$: unrelated pictures and related words). Grade 5's relearned more at 2 days ($\underline{f}_1$: unrelated words). The later retention interval (16'-30) had more relearning than the earlier one. The younger children were better at 16'-day ($\underline{f}_2$: unrelated pictures) but also better at times on 30-days ($\underline{f}_1$ and $\underline{f}_3$: unrelated words). The older children relearned more at 16'-days for unrelated words ($\underline{f}_1$) but also at 30-days for unrelated pictures ($\underline{f}_1$ and $\underline{f}_2$). The trend for the average number of consecutive errors was an increase across trials, which meant a decrease in retrieval relearning, however no significant differences were found on this measure collapsed across age, lists or retention interval.

To summarize, both storage- and retrieval-based forgetting occurred, and overall more forgetting occurred at the second interval (16'-30). Restorage and success-contingent retrieval relearning were minimal, and error-contingent retrieval relearning occurred mostly during the second interval (16'-30).

Time Effects

Comparisons were made of the first test time (Time 1 [T1]=2- v 16'-days) and of the second test time (Time 2 [T2]=16- v 30-days).

Forgetting. The greatest effect was found for storage-based forgetting, with virtually all of this forgetting being greater at 16'-days for T_1 and at 30-days for T_2. Storage-based forgetting at T_1 was higher on 16'-days than 2-days for both grades 2 and 5 on all lists with the exception of grade 5 unrelated pictures. Retrieval-based forgetting was less prominent, being higher on 16'-days than 2-days for grade 2 related words, grade 5 unrelated pictures, unrelated words and related words, and higher on 2-days than on 16'-days for grade 2 unrelated pictures. Further evidence of greater storage-based forgetting is provided by average failure rates. Differences in storage and retrieval forgetting rates were greater at 16'-days than at 2-days. At 16'-days average storage failure was .38 for grade 2's and .27 for grade 5's. The corresponding retrieval failure rate was .12 and .18 for grades 2 and 5, respectively. At 2-days the storage failure rate was .14, the retrieval rate .08 for both grades 2 and 5.

At T_2 (16- v 30-days) storage-based forgetting was greater at 30-days than 16-days for grade 2's on unrelated pictures, unrelated words and related words, and for grade 5 on related words. Retrieval-based forgetting was higher at

30-days than at 16-days for unrelated pictures, grade 2, and unrelated words, grade 5. Differences in average storage and retrieval forgetting rates were greater at 30-days than at 16-days. At 30-days average storage failure was .35 for grade 2's, .25 for grade 5's. Average retrieval failure was .06 and .13 for grades 2 and 5, respectively. At 16-days the storage failure rate was .23 and .19 for grades 2 and 5, the corresponding retrieval failure rates .04 and .09.

Relearning. Once again very little restorage and success-contingent retrieval relearning occurred. The error-contingent retrieval relearning showed no distinct trends for either T_1 or T_2.

For restorage grade 2's showed more on day 2 than on day 16' for unrelated pictures, the only difference on T_1 for either grade. The younger children had only one restorage difference on T_2 as well, related pictures greater at 16-days than at 30-days. Grade 5's had two differences on restorage, both greater at 16-days than at 30-days; unrelated pictures and related words. Neither average nor cumulative restorage rates differed between grades for T_1 or T_2. Average rates at T1 for grades 2 and 5 were .06 and .04, respectively, at T_2, .04 and .05. Cumulative rates at T_1 were .22 and .14, at T_2 .15 and .17, for grades 2 and 5, respectively.

Only one difference was found in each of T_1 and T_2 for success-contingent retrieval relearning. At T_1, grade 5's had greater relearning on day 2 than on day 16′ for unrelated pictures ($\underline{r}_1$). At T_2, grade 2's had greater relearning at day 16 than at day 30 for unrelated words ($\underline{r}_1$). Error-contingent retrieval relearning for grade 2's at T_1 was sometimes higher at 2-days ($\underline{f}_1$: unrelated words, related words), while at other times higher at 16′-days ($\underline{f}_2$: related pictures and words). For grade 5's the only difference at T_1 was greater relearning at 16′-days ($\underline{f}_2$: related pictures). At T_2, the time at which relearning was greater again varied. For grade 2's, relearning was higher at 16-days ($\underline{f}_1$: unrelated pictures; $\underline{f}_2$: unrelated pictures and words) or at 30-days ($\underline{f}_2$: related pictures; $\underline{f}_3$: related words). For both grade 5 differences, error-contingent retrieval relearning was higher at 30-days than at 16-days ($\underline{f}_1$: unrelated pictures; $\underline{f}_2$: related pictures).

To summarize, forgetting was once again the most important variable for the time comparisons (T_1= 2- $\underline{v}$ 16′-days, T_2= 16-$\underline{v}$ 30-days), particularly storage-based forgetting. The effects at relearning were less prevalent and trends were unclear.

Test and Time Effects

The last category of comparisons was a combination of test and time effects. Day 16 test, which had a preceding test, was compared with day 16' test, with no preceding test.

Forgetting. The effect of a preceding test was evident in the forgetting results. All of the significant forgetting comparisons found greater forgetting at 16'-days than at 16-days, with storage-based forgetting being the most prevalent. Storage-based forgetting for grade 2's was greater on 16'-days than on 16-days for all lists but related pictures. The only grade 5 difference was on unrelated words, again greater at 16'-days than at 16-days. Retrieval-based forgetting was also higher on 16'-days than on 16-days, for grade 2's unrelated and related words, for grade 5's unrelated pictures and related words. Storage-based forgetting was greater than retrieval-based forgetting, as measured by average failure rates. At 16'-days, storage failure was .38 and .27 for grades 2 and 5, respectively, whereas retrieval failure was .12 and .18. At 16-days, average storage failure was .23 and .19 for grades 2 and 5, respectively, retrieval failure was .04 and .09.

Relearning. Again, relearning effects were minimal. Only two restorage differences occurred, for grade 5's at 16-days on unrelated pictures and related words. The only success-contingent retrieval relearning difference favored

16-days over 16′-days for grade 5′s ($\underline{r}_1$: unrelated pictures). Error-contingent retrieval relearning was higher on 16-days than on 16′-days for the grade 2′s ($\underline{f}_1$: unrelated pictures, unrelated words and related words). For grade 5′s, error-contingent retrieval relearning was higher on 16′-days than on 16-days ($\underline{f}_1$: unrelated words; $\underline{f}_2$: related pictures).

The prior test (at 2-days) resulted in less forgetting at 16-days than if no prior test had been given (16′-days). The relearning that did occur indicated no clear trend towards either 16- or 16′-days.

To summarize the test and time comparisons, forgetting was the most prominent variable, especially storage-based forgetting, and particularly for the younger children. There was less forgetting between the first and second tests (no matter when the second test occurred, 16-days or 30-days) than between the end of acquisition and the first test, indicating that the first test affected the rate of forgetting. This is also shown with the retesting assessment with time held constant; the 16-day test, with a prior test administered, had less forgetting than the 16′-day test, with no prior testing.

DISCUSSION

Earlier, it was suggested that methodological problems, such as not equating the level of original learning at acquisition, and analytical omissions, such as not separating the forgetting and relearning components of long-term retention, led to the lack of developmental interactions in retention reported previously. These deficiencies resulted in an inadequate understanding of the development of long-term retention in children, a problem which in turn precipitated this research. Correction of the methodological and analytical problems might reveal developmental differences and the present experiment was designed to investigate this hypothesis. Indeed, when the level of original learning was equated by requiring that all subjects meet a strict acquisition criterion, and when forgetting processes (both storage- and retrieval-based) were differentiated from the processes involved in relearning (both restorage and retrieval relearning) the results of the present experiment showed clear Age X Retention interactions.

The overall findings of this experiment revealed forgetting to be the most prevalent cause of differences observed between the younger and older children's retention. As well, storage processes were found to be at least as important as retrieval, if not more so, at retention. The finding that storage processes were important at retention is significant for two reasons. First, this finding is important because theories of developmental changes at acquisition stress the importance of retrieval processes (e.g., Howe et al., 1985). In contrast to acquisition, the present results indicate that storage processes are important to the development of long-term retention. This difference provides further support for the need to study retention processes independent of acquisition. Second, the importance of storage processes at retention goes against the hypotheses that long-term retention is controlled mainly by retrieval processes.

As mentioned, an important consideration for the examination of retention performance is the separation of the forgetting and relearning components of long-term retention and within each component, the separation of storage and retrieval processes to determine the contribution of each to long-term retention. The following discussion is organized around these issues.

Forgetting

Because the mathematical model used for the analyses partitioned forgetting into storage-based and retrieval-based components, a more detailed examination of the loci of the recall failures was permitted. Of particular note was the amount of storage-based forgetting that occurred. Specifically, regardless of age, storage failure (grade 2 X=.27, grade 5 X=.21) was greater than retrieval failure (grade 2 X=.08, grade 5 X=.12). The importance of storage failure as a factor in forgetting supports the contention that retrieval processes are not the sole contributor to changes in retention performance. Rather, changes in what is in storage may also occur (see Loftus & Loftus, 1980) such that the availability of information is affected as well as the retrievability of that information. Importantly, the current results _do not_ support the trace absence view of storage failure since traces were restored after the retention interval. It would appear, then, that trace _unavailability_ should be considered as well as trace _inaccessibility_ when examining the reasons for forgetting/amnesia.

Developmental differences in forgetting can also be examined more completely than just the absolute magnitude of recall failure. Here, the grade 2 children exhibited more storage-based forgetting compared to the grade 5's, whereas the older children exhibited more retrieval-based forgetting. This age difference cannot be attributed to poorer encoding at acquisition by the younger children as the level of learning was equated across ages. Instead, differences in the type of forgetting (storage or retrieval) by age may be due to different processes used by the younger and older children to maintain traces.

The effect of modality was minimal with respect to forgetting. The few differences that were observed favored pictures over words for both grades, but neither type of forgetting, storage- or retrieval-based, was predominant. The semantically unrelated clusters were forgotten more than the related, as might be expected. If trace integrity is thought of as features bonded together to create a trace in memory, then any factor that creates and or maintains those bonds should aid the featural integrity of the trace and thus maintenance in memory (Howe & Brainerd, in press). The common category features of the related clusters provide a bond to hold the traces together and, therefore, such traces are forgotten less frequently than unrelated clusters. As well, the type of forgetting found most frequently with the unrelated clusters was storage-based. If storage-based forgetting is thought of as 'further along' the continuum than retrieval-based, it would appear that unrelated clusters are not just more difficult to retrieve but are not easily maintained intact in memory.

Storage failure was also the chief form of forgetting found as a function of the time of testing. As anticipated, greater forgetting was found on the later retention tests (16'-days and 30-days) than on the early tests

(2-days and 16-days) and this forgetting was mostly storage-based. Two retesting effects emerged. One was related to the timing of the first test. For the first test given at 2-days, forgetting increased at 16-days. When the first test was given at 16'-days, very little extra forgetting occurred at 30-days. It would appear, then, that the timing of a prior test is an important consideration when examining the effects of retesting on retention performance. The other retesting effect was related to time of test and the presence or absence of a preceding test. Greater forgetting was found at 16'-days than at 16-days due to the latter being a second test of retention. That is, subjects tested at 16-days were also tested previously at 2-days after acquisition, whereas those tested at 16'-days had not been tested before. Thus, while there was forgetting between 2 and 16 days, the early test had the effect of attenuating forgetting. The type of forgetting for this effect for grade 2's was storage-based while for the grade 5's it was retrieval-based. This finding again points to the importance of analytically separating the storage and retrieval components of forgetting when assessing developmental change as the loci of forgetting differences appear to vary with age.

The existence of restorage, discussed below, as a significant factor in recall improvements over tests trials would indicate that storage-based forgetting is not synonymous with complete absence of the memory trace, but rather, disintegration to the point that the bonds of the trace have weakened and it is indiscriminable from other more intact traces. The features of the trace would still be in memory but not as a coherent unit. Because this disintegration is gradual, redintegration of the trace is possible on retention tests. Recall attempts over successive test trials would appear to redintegrate, or 'rebond', the features together to reform the trace, so that it becomes restored (see Howe & Brainerd, in press). Trace-absence theories of memory cannot account for the possibility of a trace being restored as they contend that the trace is completely removed from memory. Restorage is consistent with the trace-integrity hypothesis of Howe and Brainerd (in press), which views both forgetting and relearning as processes related to the integrity of the bonds that form a memory trace.

The strong storage-based forgetting results found here also make it clear that forgetting is not just a trace irretrievability phenomenon, either. It would seem that current theories of memory regarding long term retention are in need of revision to include both storage and retrieval components for forgetting and retrieval.

Relearning

While it is often found that long-term retention performance improves across test trials, the reasons for this recovery (hypermnesia) are not clear. The model used here analyzed the relearning which occurred in this experiment during retention by dividing it into two components, restorage and retrieval relearning, in order to differentiate the reasons for the occurrence of hypermnesia. From these analyses several findings appeared.

Restorage was a central factor responsible for the increase in net recall, hypermnesia, across test trials. This was the case across all variables; that is, restorage did not vary consistently relative to differences in age, presentation modality, semantic relatedness, or the number or timing of retention tests. It would seem then that redintegration of a trace is possible with more than just semantically related information (see Howe & Brainerd, in press). The importance of restorage to hypermnesia is consistent with the disintegration/redintegration hypothesis. Retrieval relearning after an success (r's) also contributed to hypermnesia to some extent. It was mainly constant but at times had a slight tendency to increase across trials. That is, the probability of successful recall after a correct response increased as the number of consecutive successes increased. Retrieval relearning after an error (f's) deteriorated across test trials. That is, the probability of successful recall after an error declined as the number of consecutive errors increased and therefore was not a factor in hypermnesia. As with restorage, few differences were found in retrieval relearning between ages, the semantic or modality manipulations.

Hypermnesia, as measured by restorage and retrieval relearning, was not always dependent on semantic relatedness or the mode of presentation. This runs counter to previous literature which found such differences with adults (e.g., Erdelyi, Buschke, & Finkelstein, 1977). Further, the length of the retention interval or the number of tests did not affect the net increase in recall over tests within any of the testing sessions (although ceiling effects have to be considered as a factor for some of the comparisons). Hypermnesia, then would seem to be a result of repeated testing (Howe & Brainerd, in press; Payne, 1987).

Conclusions

It appears that developmental trends do exist, particularly with forgetting, in children's long-term retention. This runs counter to previous research and general opinion (e.g., Lehman et al, 1985). Controlling such variables as the level of learning at acquisition and the separating of forgetting and relearning may have unmasked previously hidden trends.

As stated at the outset, the understanding of memory development will come from delineation of the variables that affect both the acquisition and retention of information. The present finding that both storage and retrieval were involved in amnesia and hypermnesia indicates that the argument over trace retrievability _or_ trace accessibility should go the way of the nature/nurture argument. Instead, it would be more valuable to delineate the conditions under which amnesia and hypermnesia occur and the variables that affect storage and retrieval processes at retention. Both storage and retrieval are important in long-term retention and are likely different components of the same phenomenon. The disintegration/redintegration hypothesis of Howe and Brainerd (in press) is based on this assumption and can be incorporated here to account for the present findings. Furthermore, this hypothesis integrates amnesia and hypermnesia so that they can also be viewed as two components of the same phenomenon, namely, trace integrity. That is, the disintegration of the trace is related to amnesia while the redintegration of the trace is related to hypermnesia.

Finally, the benefits of using a mathematical model with independent parameters to differentiate and assess the contributions of the components of long-term retention are clear. These analytical refinements, coupled with the methodological improvements instituted in this research, permitted the previously obscured developmental trends in children's long-term retention to be observed and evaluated in a theoretical framework.

Table 1

Theoretical Definitions of the Retention Model's Parameters

Process and Parameter	Theoretical Definition
Forgetting	
$\underline{S}$	The probability of storage failure.
$\underline{R}$	For information that is in storage (or is subsequently restored), the probability of retrieval failure.
Relearning	
$\underline{a}$	For information not in storage, the probability of restorage on any test trial.
$\underline{r}_1$	The probability that stored (or restored) information is successfully recalled following a success on the immediately preceding trial.
$\underline{r}_2$	The probability that stored (or restored) information is successfully recalled following successes on the two immediately preceding trials.
$\underline{r}_3$	The probability that stored (or restored) information is successfully recalled following successes on the three immediately preceding trials.

Table 1 (cont'd)

$\underline{f}_1$	The probability that stored (or restored) information is successfully recalled following an error on the immediately preceding trial.
$\underline{f}_2$	The probability that stored (or restored) information is successfully recalled following errors on the two immediately preceding trials.
$\underline{f}_3$	The probability that stored (or restored) information is successfully recalled following errors on the three immediately preceding trials.

Table 2

The Retention Model's Theoretical Expressions for the 16 Probabilities in the Empirical Outcome Space

Outcome probability	Theoretical Expression
$p(CCCC)$	$(1-S)(1-R)r_1r_2r_3$
$p(CCCE)$	$(1-S)(1-R)r_1r_2(1-r_3)$
$p(CCEC)$	$(1-S)(1-R)r_1(1-r_2)f_1$
$p(CECC)$	$(1-S)(1-R)(1-r_1)f_1r_1$
$p(ECCC)$	$Sa(1-R)r_1r_2 + (1-S)Rf_1r_1r_2$
$p(CCEE)$	$(1-S)(1-R)r_1(1-r_2)(1-f_1)$
$p(CECE)$	$(1-S)(1-R)(1-r_1)f_1(1-r_1)$
$p(ECCE)$	$Sa(1-R)r_1(1-r_2) + (1-S)Rf_1r_1(1-r_2)$
$p(CEEC)$	$(1-S)(1-R)(1-r_1)(1-f_1)f_2$
$p(ECEC)$	$Sa(1-R)(1-r_1)f_1 + (1-S)R\imath_1(1-r_1)f_1$
$p(EECC)$	$S(1-a)a(1-R)r_1 + SaRf_1r_1 + (1-S)R(1-f_1)f_2r_1$
$p(CEEE)$	$(1-S)(1-R)(1-r_1)(1-f_1)(1-f_2)$
$p(ECEE)$	$Sa(1-R)(1-r_1)(1-f_1) + (1-S)Rf_1(1-r_1)(1-f_1)$
$p(EECE)$	$S(1-a)a(1-R)(1-r_1) + SaRf_1(1-r_1) +$ $(1-S)R(1-f_1)f_2(1-r_1)$
$p(EEEC)$	$S(1-a)^2a(1-R) + S(1-a)aRf_1 + SaR(1-f_1)f_2 +$ $(1-S)R(1-f_1)(1-f_2)f_3$
$p(EEEE)$	$S(1-a)^3 + S(1-a)^2aR + S(1-a)aR(1-f_1) +$ $SaR(1-f_1)(1-f_2) + (1-S)R(1-f_1)(1-f_2)(1-f_3)$

Note: C = correct response E = incorrect response

Table 3

Mean Number of Correctly Recalled Clusters

Grade and List	LTR 1				LTR 2			
	1	2	3	4	1	2	3	4
Grade 2								
Unrelated pictures								
2-16 day	3.75	4.15	4.00	4.35	3.20	3.35	3.35	3.85
16'-30 day	2.05	2.20	2.30	2.75	2.35	2.65	2.55	2.80
Related pictures								
2-16 day	5.75	5.80	5.80	5.80	5.65	5.60	5.65	5.70
16'-30 day	5.00	5.20	5.35	5.30	5.30	5.35	5.45	5.50
Unrelated words								
2-16 day	3.65	3.70	4.10	4.05	3.20	3.50	3.55	3.55
16'-30 day	1.85	2.10	2.10	2.20	2.30	2.10	2.35	2.35
Related words								
2-16 day	5.85	5.85	5.85	5.90	5.55	5.70	5.75	5.80
16'-30 day	4.15	4.20	4.70	4.75	4.80	4.70	4.80	5.00

Table 3 (cont'd)

Grade 5								
Unrelated pictures								
2-16 day	3.95	4.00	4.10	4.45	3.35	3.70	3.95	4.15
16'-30 day	2.85	2.75	2.85	2.90	3.00	3.35	3.30	3.25
Related pictures								
2-16 day	5.80	5.95	5.95	5.95	5.75	5.85	5.85	5.90
16'-30 day	4.95	5.45	5.40	5.50	5.50	5.50	5.60	5.55
Unrelated words								
2-16 day	3.65	3.75	3.85	4.10	3.45	3.45	3.50	3.65
16'-30 day	2.50	2.55	3.00	3.00	2.80	2.85	2.75	2.70
Related words								
2-16 day	5.70	5.85	5.85	5.85	5.40	5.45	5.60	5.65
16'-30 day	4.65	4.90	4.90	5.00	4.70	4.90	4.80	4.85

Note: Columnar values out of a possible 6.00

Table 4

Goodness of Fit Assessment of the Long-Term Retention Model

Grade and list condition		$-2\ln L_9$	$-2\ln L_{15}$	$X^2(6)$
Grade 2				
Unrelated pictures	2-days	396.88	384.31	12.57
	16-days	389.04	387.38	1.66
	16'-days	357.28	347.97	9.31
	30-days	341.99	333.62	8.37
Related pictures	2-days	46.62	46.57	0.05
	16-days	101.27	98.79	2.48
	16'-days	174.49	174.06	0.43
	30-days	130.97	130.31	0.66
Unrelated words	2-days	323.82	320.45	3.37
	16-days	294.97	280.60	14.37
	16'-days	265.99	249.67	16.32
	30-days	291.06	282.74	8.32
Related words	2-days	31.92	31.87	0.05
	16-days	105.03	104.10	0.93
	16'-days	309.39	304.12	5.27
	30-days	228.61	222.83	5.78

Table 4 (cont'd)

Grade 5				
Unrelated pictures	2-days	345.37	345.21	0.16
	16-days	379.43	375.16	4.27
	16'-days	400.04	392.71	7.33
	30-days	369.00	365.57	3.43
Related pictures	2-days	39.62	39.57	0.05
	16-days	54.74	52.11	2.63
	16'-days	191.70	181.50	10.20
	30-days	102.84	93.01	9.83
Unrelated words	2-days	304.96	301.58	3.38
	16-days	357.48	352.59	4.89
	16'-days	351.95	344.65	7.30
	30-days	290.96	284.64	6.32
Related words	2-days	56.03	55.96	0.07
	16-days	168.72	159.72	9.00
	16'-days	220.32	215.96	4.36
	30-days	173.07	172.98	0.09

Note: For goodness-of-fit the value of $X^2(6)$ must not be greater than 16.81 ($p < .01$).

Table 5

Estimates of the Retention Model's Theoretical Parameters

Grade and List	S	R	a	r_1	r_2	r_3	f_1	f_2	f_3
Grade 2									
Unrelated pictures									
2-days	.25	.14	.22	.86	.93	.99	.48	.10	.34
16-days	.44	.00	.11	.83	.95	.98	.79	.50	.34
16'-days	.65	.00	.08	.87	.93	1.0	.55	.34	.41
30-days	.54	.11	.05	.87	.94	1.0	.59	.00	.31
Related pictures									
2-days	.03	.01	.03	1.0	1.0	1.0	.95	.00	.00
16-days	.05	.01	.10	.98	.99	1.0	.67	.00	.00
16'-days	.14	.02	.10	.99	1.0	.98	.96	.97	.15
30-days	.08	.04	.00	.98	1.0	1.0	.66	.97	.00
Unrelated words									
2-days	.27	.17	.02	.93	1.0	.96	.62	.39	.35
16-days	.39	.12	.03	.96	.96	1.0	.67	.31	.46
16'-days	.60	.24	.02	.96	.95	1.0	.33	.22	.12
30-days	.61	.03	.03	.85	.95	.97	.85	.00	.96
Related words									
2-days	.01	.01	.00	1.0	1.0	1.0	.70	.00	.29
16-days	.03	.04	.00	.99	.99	1.0	.86	.39	.10
16'-days	.12	.21	.00	.95	.99	.97	.32	.51	.13
30-days	.16	.05	.01	.94	.99	1.0	.55	.39	.63

Table 5 (cont'd)

Grade 5									
Unrelated pictures									
2-days	.22	.15	.02	.90	.97	.99	.43	.46	.41
16-days	.31	.18	.10	.89	.97	1.0	.47	.37	.00
16'-days	.29	.33	.00	.78	.96	.96	.44	.27	.19
30-days	.37	.19	.00	.92	.89	.94	.70	.48	.24
Related pictures									
2-days	.01	.03	.00	1.0	1.0	1.0	.69	.00	.00
16-days	.03	.00	.09	.98	1.0	1.0	.68	.15	.00
16'-days	.11	.06	.08	.98	.99	1.0	.60	.82	.23
30-days	.07	.02	.00	.98	1.0	1.0	.66	.95	.00
Unrelated words									
2-days	.30	.10	.09	.92	1.0	.97	.61	.28	.41
16-days	.31	.17	.10	.89	.97	.97	.27	.20	.13
16'-days	.58	.20	.08	.87	.96	1.0	.62	1.0	.46
30-days	.36	.27	.00	.89	.96	.98	.25	.00	.07
Related words									
2-days	.03	.03	.00	1.0	1.0	1.0	.68	.00	.00
16-days	.09	.01	.12	.96	.99	1.0	.60	.00	.00
16'-days	.11	.12	.00	.96	.99	1.0	.60	.00	.19
30-days	.18	.03	.00	.99	.99	.99	.68	.01	.04

Table 6

Conditionwise Tests

Effects	2-day	16-day	16'-day	30-day
Developmental Effects				
Unrelated pictures	22.48	26.60	24.70	27.37
Related pictures	ns	ns	ns	ns
Unrelated words	ns	23.45	30.38	25.60
Related words	32.90	29.11	25.87	ns
Modality Effects				
Grade 2 unrelated	42.27	21.94	ns	ns
related	ns	ns	54.63	27.06
Grade 5 unrelated	ns	ns	22.56	28.86
related	ns	ns	22.48	23.76
Semantic Effects				
Grade 2 pictures	256.25	191.94	213.91	214.51
words	307.00	196.74	124.09	121.54
Grade 5 pictures	210.81	203.71	182.84	171.76
words	186.53	137.53	129.83	94.99

Table 6 (cont'd)

	Unrelated pictures	Related pictures	Unrelated words	Related words
T^1 v T^2 Effects				
Grade 2 2 v 16 day	32.01	ns	21.98	63.28
16' v 30 day	22.65	ns	25.53	22.65
Grade 5 2 v 16 day	ns	ns	27.66	33.33
16' v 30 day	32.94	ns	21.88	ns
16 v 16' Effects				
Grade 2	28.01	ns	32.90	110.85
Grade 5	29.59	47.73	22.38	39.05
T^1 (2 v 16') Effects				
Grade 2	82.88	38.03	65.66	258.97
Grade 5	35.79	59.20	41.35	45.42
T^2 (16 v 30) Effects				
Grade 2	21.95	21.83	34.88	67.84
Grade 5	35.85	22.05	22.50	56.40

Note: Columnar values are $X^2(9)$ $p < .01$ significant at 21.67.

ns = not significant

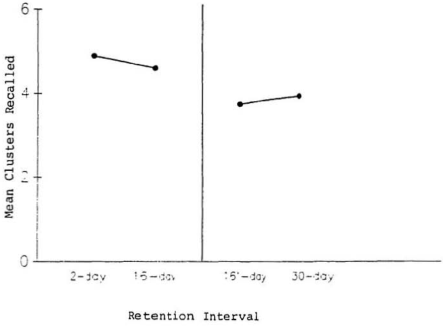

Figure 1: Mean number of clusters recalled across the different retention test sessions (collapsed across ages, list conditions, and trials).

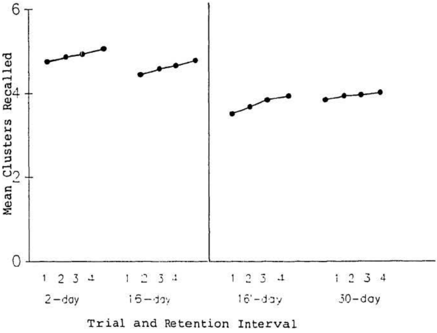

Figure 2: Mean number of clusters recalled for each test trial across the different retention test sessions (collapsed across ages and list conditions).

References

Ammons, H., & Irion, A.L. (1954). A note on the Ballard reminiscence phenomenon. Journal of Experimental Psychology, 48, 184-6.

Ballard, P.B. (1913). Obliviscence and reminiscence. British Journal of Psychology Monograph Supplements, 1, 1-82.

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

Belli, R.F. (1989) Influences of misleading postevent information: Misinformation interference and acceptance. Journal of Experimental Psychology: General, 118, 72-85.

Belmore, S.M. (1981). Imagery and semantic elaboration in hypermnesia for words. Journal of Experimental Psychology: Human Learning and Memory, 7, 191-203.

Brainerd, C.J., Howe, M.L., & Kingma, J. (1982). An identifiable model of two-stage learning. Journal of Mathematical Psychology, 26, 263-293.

Brainerd, C.J., Kingma, J. & Howe, M.L. (1985). On the development of forgetting. Child Development, 56, 1103-1119.

Brainerd, C.J. & Reyna, V.F. (1988). Memory loci of suggestibility development: Comment on Ceci, Ross, and Toglia (1987). Journal of Experimental Psychology: General,117, 197-200.

Brainerd, C.J. & Reyna, V.F. (1989). Development of retention: Amnesia, hypermnesia, forgetting, and reminiscence. Manuscript submitted for publication.

Brainerd, C.J. & Reyna, V.F. (in press). The gist is the grist: Fuzzy-trace theory and the new intuitionism. Developmental Review.

Ceci, S.J., Ross, D.F., & Toglia, M.P. (1987). Suggestibility in children's memory: Psycholegal implications. Journal of Experimental Psychology: General, 116, 38-49.

Ceci, S.J., Toglia, M.P., & Ross, D.F. (1988). On remembering ... more or less: A trace strength interpretation of developmental differences in suggestibility. Journal of Experimental Psychology: General, 117, 201-203.

Erdelyi, M.H., & Becker, J. (1974). Hypermnesia for pictures: Incremental memory for pictures but not for words in multiple recall trials. Cognitive Psychology, 6, 159-171.

Erdelyi, M.H., Buschke, H., & Finkelstein, S. (1977). Hypermnesia for socratic stimuli: The growth of recall for an internally generated memory list abstracted from a series of riddles. Memory & Cognition, 5, 283-286.

Fajnsztejn-Pollack, G. (1973). A developmental study of decay rate in long-term memory. Journal of Experimental Child Psychology, 16, 225-35.

Hasher, L., & Thomas, H. (1973). A developmental study of retention. Developmental Psychology, 9, 281.

Horowitz, L.M., & Prytulak, L.S. (1969). Redintegrative memory. Psychological Review, 76, 519-531.

Howe, M.L. (1985). Storage and retrieval of associative clusters: A stages-of-learning analysis of associative memory traces. Canadian Journal of Psychology, 39, 34-53.

Howe, M.L. (1987). The development of forgetting in childhood. Paper presented at the biennial meeting of the Society for Research in Child Development, Baltimore, Maryland.

Howe, M.L. & Brainerd, C.J. (in press). Development of children's long-term retention. Developmental Review.

Howe, M.L., Brainerd, C.J., & Kingma, J. (1985). Development of organization in recall: A stages-of-learning analysis. Journal of Experimental Child Psychology, 39, 230-251.

Howe, M.L., & Hunter, M.A. (1986). Long-term memory in adulthood: An examination of the development of storage/retrieval processes at acquisition and retention. Developmental Review, 6, 334-364.

Hudson, J., & Nelson, K. (1986). Repeated encounters of a similar kind: Effects of familiarity on children's autobiographic memory. Cognitive Development, 1, 253-271.

Lehman, E.B., Mikesell, J.W., & Doherty, S.C. (1985). Long-term retention of information about presentation modality by children and adults. Memory & Cognition, 13, 21-28.

Loftus, E.F., Hoffman, H.G. & Wagenaar, W.A. (in press). The misinformation effect: Transformations in memory induced by postevent information. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), Development of long-term retention. New York: Academic Press.

Loftus, E.F. & Loftus, G.R. (1980). On the permanence of stored information in the human brain. American Psychologist, 35, 409-20.

McCloskey, M. & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. Journal of Experimental Psychology: General, 114, 1-16.

Paris, S.G. (1978). Memory organization during children's repeated recall. Developmental Psychology, 14, 99-106.

Payne, D. G. (1987). Hypermnesia and reminiscence in recall: A historical and empirical review. Psychological Bulletin, 101, 5-27.

Penfield, W. (1969). Consciousness, memory, and man's conditioned reflexes. In K. Pribram (Ed.), On the biology of learning. New York: Harcourt, Brace & World.

Piaget, J. & Inhelder, B. (1973). Memory and Intelligence. New York: Basic Books.

Richardson, J.T.E. (1985). The effects of retention tests upon human learning and memory: A historical review and an experimental analysis. Educational Psychology, 5, 85-114.

Roediger, H.L., Payne, D.G. (1982) Hypermnesia: the role of repeated testing. Journal of Experimental Psychology: Learning Memory and Cognition, 8, 66-72.

Rogoff, B., Newcombe, N., & Kagan, J. (1974). Planfulness and recognition memory. Child Development, 45, 872-877.

Runquist, W.N. (1986a). The effect of testing on the forgetting of related and unrelated associates. Canadian Journal of Psychology, 40, 65-76.

Runquist, W.N. (1986b). Changes in the rate of forgetting produced by recall tests. Canadian Journal of Psychology, 40, 282-289.

Runquist, W.N. (1987). Retrieval specificity and the attenuation of forgetting by testing. Canadian Journal of Psychology,41, 84-90.

Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. Journal of Experimental Psychology: Human Learning and Memory, 6, 174-215.

Sophian, C., & Perlmutter, M. (1980). Encoding and retention factors in the early development of recall. Bulletin of the Psychonomic Society, 15, 342-344.

Toglia, M.P., & Battig, W.F. (1978). Handbook of semantic word norms. Hillsdale, N.J.: Erlbaum.

Tversky, B. & Tuchin, M. A reconciliation of the evidence on eyewitness testimony: Comments on McCloskey and Zaragoza. Journal of Experimental Psychology: General, 118, 86-91.

Vertes, J.O. (1931/32). Behalten und vergessen des kindes.(Retention and forgetting in the child.) Zsch. f.Psychologie, 122, 241-354. Psychological Abstracts, 1932, 1221, 145.

Zaragoza, M.S. & McCloskey, M. Misleading postevent information and the memory impairment hypothesis: Comment on Belli and reply to Tversky and Tuchin. Journal of Experimental Psychology: General, 118, 92-99.

Appendix A

Unrelated Clusters	Related Clusters
Train: Pie - Ear	Train: Bus - Plane
Sock: Banana - Cup	Sock: Coat - Pants
Apple: Coat - Horse	Apple: Banana - Grapes
Bread: Cow - Pants	Bread: Pie - Cake
Glass: Arm - Grapes	Glass: Bowl - Cup
Pig: Bus - Cake	Pig: Cow - Horse

Appendix B

The unrestricted likelihood of the recall data was calculated for each of the 32 Grade x LTR session x Semantic x Modality x Retention interval combinations. There are a total of 16 possible outcomes in the data space (e.g., CCCC, CCCE, ..., EEEE). Probabilities can be attached to each of these events [e.g., p(CCCC), p(CCCE), ..., p(EEEE)]. According to the theory of maximum likelihood, a function can be written which gives the a posteriori probability, or likelihood, of a sample data set:

$$L_15 = [\underline{p}(CCCC)]^{N(CCCC)} \times [\underline{p}(CCCE)]^{N(CCCE)} \times \ldots \times [\underline{p}(EEEE)]^{N(EEEE)}$$

The exponents represent the frequency of occurrence of each of the events in the outcome space. This function has 15 degrees of freedom (parameters) and the goodness of fit of any model with fewer then 15 parameters can be evaluated by comparing the aposteriori likelihood of the same data under the model's assumptions. For the 9-parameter model, the theoretical expressions of the model are replaced by the probability terms in Table 2 (e.g., [p(CCCC)] written as $[(1-\underline{S})(1-\underline{R})\underline{r}_1\underline{r}_2\underline{r}_3])$. This probabilities function has 9 degrees of freedom and provides an estimate of the likelihood of the data, L_9. The parameter space of the long-term retention model is a portion of the empirical probability space and therefore it is known that the statistic

$$-2\ln (L_9/L_{15})$$

has an asymptotic X^2 distribution with 15-9=6 degrees of freedom. The -2ln value of each likelihood function is what is actually calculated, and this statistic is computed by simply subtracting $-2\ln L_9$ by $-2\ln L_{15}$. The likelihood that the data obtained using the long-term retention model does not differ reliably from the actual data (the null hypothesis) can be tested using this statistic.