FUZZY-TRACE THEORY: AN INTERIM SYNTHESIS

V. F. REYNA AND C. J. BRAINERD

UNIVERSITY OF ARIZONA

ABSTRACT: We review the current status of fuzzy-trace theory. The presentation is organized around five topics. First, theoretical ideas that immediately preceded the development of fuzzy-trace theory are sketched. Second, experimental findings that challenged those ideas (e.g., memory-reasoning independence, the intuitive nature of mature reasoning) are summarized. Third, the core assumptions that comprised the initial version of fuzzy-trace theory are described. Fourth, some modifications to those assumptions are explored that were necessitated by subsequent experimental findings. Fifth, four areas of experimentation are considered in which research under the aegis of fuzzy-trace theory is in progress: (a) suggestibility and false memories; (b) judgment and decision making; (c) the development of forgetting; and (d) the development of retrieval.

During the past few years, we have been formulating an approach to cognitive development, fuzzy-trace theory, that began as an attempt to extract theoretical meaning from some experimental results that appeared in the literature from the mid-1980s to the early 1990s. Those results were of three general kinds: (a) findings that underscored the noncomputational, intuitive character of sophisticated forms of human reasoning (e.g., deductive inference); (b) findings that demonstrated independence between such reasoning and accurate memory for critical informational inputs to reasoning; and (c) findings that demonstrated surprising dissociations between memory for different aspects of the same informational inputs, particularly between memory for their precise surface forms and memory for their patterns and meanings.

To make sense of such findings, it has been necessary to initiate programs of experimentation in diverse areas. Examples include attitude change, classification, discourse processes, metaphorical interpretation, dual-task interference, forgetting from long-term memory, judgment and decision making, mathematical and scientific problem solving, recall, recognition, sentential inference, short-term memory

Direct all correspondence to: V. F. Reyna, Department of Educational Psychology, University of Arizona, Tucson, AZ 85721.

Learning and Individual Differences, Volume 7, Number 1, 1995, pages 1–75.

ISSN: 1041-6080

span, and suggestibility. Such diversity requires, however, periodic efforts at theoretical stock-taking. Ideally, such stock-taking should result in identifying new areas of application, as well as boundary conditions that limit generalization; refining, changing, or eliminating theoretical assumptions in deference to data; increasing the number and specificity of predictions, eventually leading to formal models; and integrating the implications of new research into a theoretically coherent view of the domain of inquiry. The purpose of this article is to move toward these goals.

The discussion proceeds in five steps. First, we sketch the general situation in cognitive-developmental theory as of about 15 years ago. We examine what, at that time, were the two standard hypotheses about the relationship between memory and cognitive development, hypotheses that have been repeatedly disconfirmed in experimentation: *memory necessity* (accurate memory is a precondition for good reasoning) and *constructivism* (memory is shaped by reasoning). Second, we consider the types of experimental results that originally led us to conclude that there were fundamental difficulties with then-dominant theories of cognitive development and that an alternative framework was required to explain them. Findings of independence between the accuracy of children's reasoning (e.g., on transitive inference or probability judgment problems) and the accuracy of their memory for critical informational inputs are emphasized, as are findings showing that mature reasoning (e.g., decision making) tends to process the gist of inputs rather than memories of their exact content. Third, we consider fuzzy-trace theory's core assumptions about reasoning, memory, and the relationships between them. This is the initial version of the theory, our first pass at accounting for the results that are summarized in the second section of the article.

Fourth, we discuss some major adjustments to the initial version of the theory that were made to accommodate new findings. Those adjustments include spelling out the conditions under which children's reasoning does not require accurate memory for critical information, introducing the notion of memory interference, formulating the concept of parallelism in the storage of verbatim and gist memories of inputs, identifying the particular gist memories that are the basis of accurate reasoning on particular tasks (e.g., class inclusion), and charting developmental changes in reliance on verbatim and gist memories. Fifth, we review research in four areas that are the loci of new developments in fuzzy-trace theory. The specific areas are suggestibility and false memories, judgment and decision making, the development of forgetting, and the development of retrieval.

Although our focus throughout is on cognitive development, much that is said has implications for the study of individual differences in children and adults. At a general level, as noted elsewhere (Reyna & Brainerd 1991a), developmental differences are special cases of individual differences. Intersubject variability is the key idea in both instances. That variability simply arises from a particular factor (age) in one case and from a host of factors (e.g., ethnicity, gender, socioeconomic status) in the other case. At a more specific level, many of the theoretical principles that we present (e.g., fuzzy-processing preferences in reasoning, sensitivity to interference in reasoning and memory, gist extraction in memory) specify performance variables that are apt to exhibit stable individual differences. Indeed, progress has been made in identifying individual differences in some of these variables (Brainerd & Reyna 1991a, in press b; Brainerd, Reyna, Harnishfeger, & Howe 1993; Harnishfeger & Brainerd 1994; Reyna & Brainerd 1991a).

HISTORICAL BACKGROUND: MEMORY-REASONING RELATIONSHIPS

Until recently, theoretical conceptions of the relationship between memory and reasoning were dominated by two ideas. The more widespread idea was the *necessity* assumption. Theorists assumed that memory was necessary for reasoning: Accurate reasoning required, at the very least, that a veridical record of the problem information be retained until it could be processed. In Piagetian paradigms, such as conservation, this seemed a trivial matter. Therefore, variations in memory for problem facts were assumed to contribute little to variations in reasoning. From the Piagetian perspective, memory variations were little more than artifacts (e.g., Smedslund 1969).

As information-processing theories gained influence, however, greater emphasis was placed on memory. Information-processing theorists used memory variations to explain performance on tests of inference (e.g., Bryant & Trabasso 1971), individual differences in reading comprehension (e.g., Daneman & Carpenter 1980), errors in decision making (e.g., Simon 1956), and performance on a host of other complex cognitive tasks (see Bjorklund 1989). The common denominator across these accounts was the idea that memory capacity (or mental resources) influenced the expression of higher cognitive abilities. This idea was compatible with certain earlier traditions, such as the assessment of memory span on intelligence tests, which were then given information-processing interpretations (e.g., Case 1985).

Thus, Piagetian scholars focused on nonmemorial explanations for reasoning, while acknowledging the necessity of memory, whereas information-processing theorists viewed memory as crucial for reasoning, while emphasizing its functional limitations (e.g., Miller 1956). It would be a mistake, however, to leave the impression that Piagetian theory had nothing to say about memory beyond necessity.

In fact, the reasoning-remembering relationship is quite complex in Piagetian theory. The general stance that Piaget advocated, called *constructivism*, held that memory is subordinated to reasoning, as captured in the well-known quote,"The schemata used by the memory are borrowed from the intelligence" (Piaget & Inhelder 1973, p. 382). Because memory is subordinated to reasoning, reasoning shapes the content of memory (Paris & Lindauer 1977; Perner & Mansbridge 1983; Prawatt & Cancelli 1976).

Constructivism was also supported, within adult psychology, by Bartlett's (1932) reports of culturally biased distortions in memory for folk tales. However, as psychological methods became more rigorous, more convincing proof of constructive memory was sought. In the 1970s, such proof seemed at hand. Bransford and colleagues published a series of experiments on sentence memory in which adults failed to discriminate sentences that had actually been presented from ones that meant the same thing (e.g., Bransford & Franks 1971; Bransford, Barclay, & Franks 1972). Paris and Carter (1973) replicated these results with children. These false-recognition effects seemed to conclusively demonstrate that what adults and children remembered was what they understood, rather than actual experience.

Thus, in the theoretical environment that immediately predated fuzzy-trace theory, close connections between reasoning and memory were assumed. As is the case with most ideas that are taken to be obvious, as well as overdetermined, researchers failed to actually test these assumptions. Logic itself seemed to require that reasoning must depend on memory for the problem. Also, demonstrations of apparently constructive memory had become conventional wisdom, and were routinely cited in textbooks (e.g., Bjorklund 1989; Schwartz & Reisberg 1991). Furthermore, the study of memory development was increasingly concerned with the effects of conceptual and strategic factors, such as subjective organization, as opposed to memory in the "strict sense" (Siegler 1991; Schneider & Pressley 1989). Memory and higher cognition had become so intertwined that it was difficult to differentiate them (Siegler 1991).

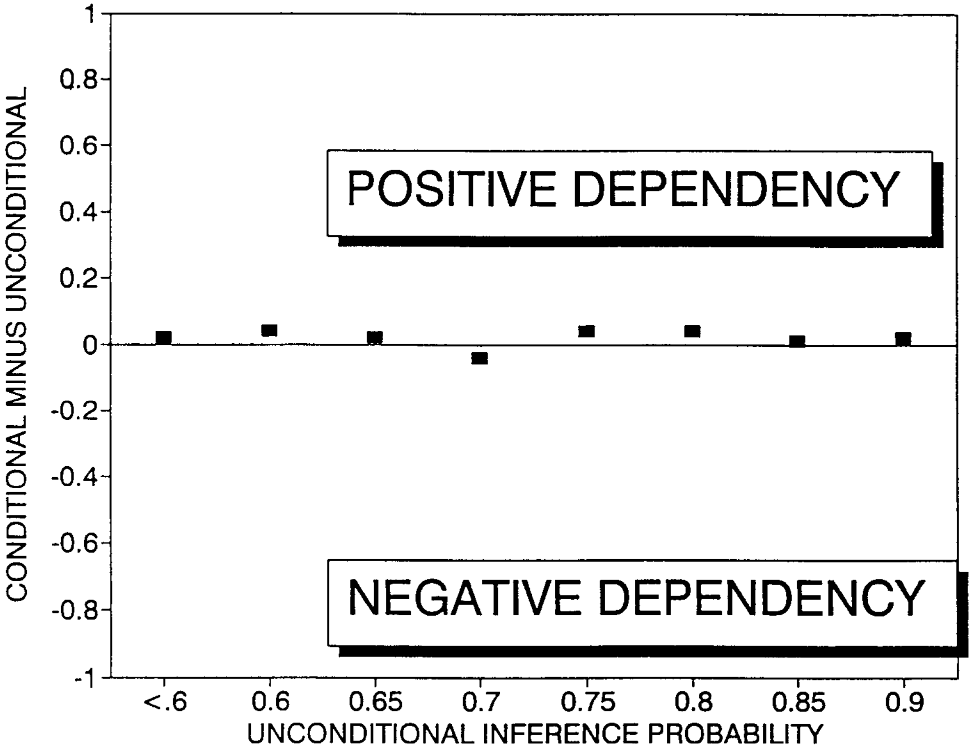
EMPIRICAL CHALLENGES

THE MEMORY-INDEPENDENCE EFFECT

In 1984, Brainerd and Kingma published a series of experiments with children using a simple modification of the standard Piagetian transitive inference task. The standard task involves presenting children with items that vary imperceptibly in magnitude along some dimension, for example, in length. Children are told about the relationships between adjacent pairs of items, for example, "The red stick is longer than the white stick; The white stick is longer than the blue stick," and so on. After presentation of these "premises," the child is asked an inference question about nonadjacent items, such as whether the red or blue stick is longer.

Brainerd and Kingma (1984) modified the standard procedure by adding memory questions about the adjacent relationships. They then tested whether success on the inference question was in any way related to success on the memory question—i.e., they measured the stochastic dependency between these responses. In eight experiments with five- to eight-year-olds, they found that reasoning performance was entirely independent of memory performance. Shortly thereafter, they reported the same independence result for probability judgment, class inclusion, and various conservation tasks (Brainerd & Kingma 1985). Data that illustrate the memory-independence result for transitive inference are shown in Figure 1, and data that illustrate the same result for the other three reasoning tasks are shown in Figure 2. In each case, the unconditional probability of correct reasoning has been plotted against the conditional probability of correct reasoning when memory

FIGURE 1
Relationship between unconditional probabilities of correct transitive inferences and conditional probabilities of correct inferences given correct memory responses. These data are from studies reviewed by Brainerd and Reyna (1992a).

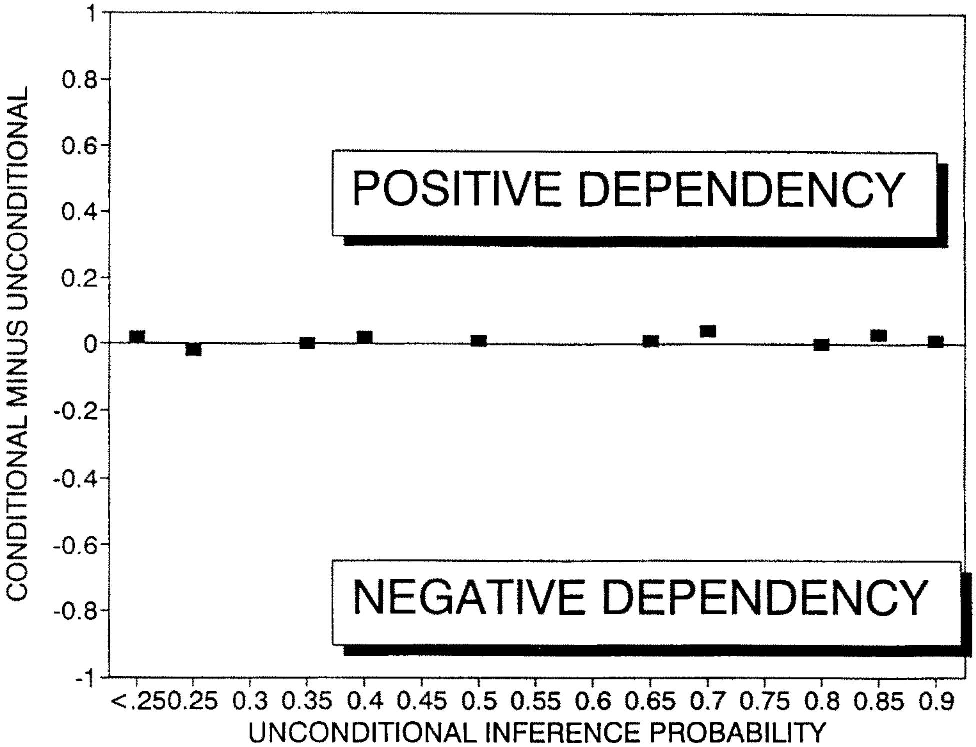


probes have been correctly answered. As can be seen, all the plotted values fall very near the zero-dependency line.

Clearly, memory-independence findings contradict predictions of necessity. If memory were necessary for reasoning, then the conditional probability of successful reasoning—reasoning performance *given* that memory performance was successful—should exceed the unconditional probability of successful reasoning. In other words, necessity implies positive dependency between memory and reasoning. If necessity obtains, positive dependency should hold, even if factors other than memory are major contributors to reasoning performance. The tests of dependency administered by Brainerd and Kingma (1984, 1985) merely assess whether any of the variation in reasoning is associated with variation in memory.

Brainerd and Kingma (1984) pointed out that then-extant theories, mostly of the information-processing type, predicted that necessity should hold for transitive reasoning, which was disconfirmed. As an alternative, they proposed that subjects used some sort of streamlined representation, a "fuzzy trace," rather than memory

FIGURE 2
Relationship between unconditional probabilities of correct inferences on class-inclusion, conservation, and probability judgment tasks and conditional probabilities of correct inferences given correct memory responses. These data are from studies reported by Brainerd and Kingma (1985).



for the actual problem facts. In their 1985 paper, they addressed predictions of another class of theories that predict memory-reasoning relationships, resource theories (Baddeley 1986). In such theories, this relationship is a by-product of the assumption that a limited-capacity processing resource constrains both memory and reasoning. Again, Brainerd and Kingma invoked the concept of fuzzy traces to explain how reasoners might have avoided relying on memory for specific problem facts.

FURTHER EVIDENCE

New data and reanalyses of old data continued to confirm the memory-independence effect (Brainerd & Reyna 1988a, 1992b; Reyna & Brainerd 1990). Such research also helped to identify boundary conditions for the phenomenon. For example, Brainerd and Reyna (1988a) found that positive dependency was obtained for tasks in which reasoning as well as memory questions required precise responses, as in mental arithmetic. Even in mental arithmetic, however, it turned out that dependency did not occur because memory was necessary for reasoning. Instead, the same kinds of reasoning processes involved in solving arithmetic problems were used to answer memory questions. Children *computed* their answers to memory questions by working backward from solutions to the arithmetic problems. For example, if a child had added 2 to 4 to obtain a sum of 6, they simply undid that operation (by subtracting 2) to recover 4 as the original addend.

It should be stressed that such findings did not support either the notion of memory necessity nor the claim that memory and reasoning are both constrained by a limited-capacity processing resource. Moreover, the finding that subjects reconstructed responses to memory questions in mental arithmetic (rather than retrieving them) contradicted the familiar assumption that immediate memory probes are answered by retrieval from short-term memory, whereas reconstruction is characteristic of long-term memory (Bartlett 1932; Kolodner 1983). Thus, data on memory-reasoning relationships also challenged two of the criterial distinctions between short- and long-term memory, namely capacity limitations and retrieval-as-readout (STM) versus unlimited capacity and retrieval-as-reconstruction (LTM).

Despite memory-independence findings, information-processing models, particularly those that exploited the idea of limited resources, seemed to account for a broad range of data. However, much of this evidence was correlational. For instance, measures of certain types of memory span were shown to correlate with individual differences in reasoning (e.g., Daneman & Carpenter 1980). As in the case of mental arithmetic, however, purported measures of "memory" also tapped reasoning processes (Turner & Engle 1989). A related finding was that age improvements in reasoning were concentrated within age ranges in which memory improved dramatically (Bjorklund 1990; Schneider & Pressley 1989). This led to the hypothesis that increases in the capacity or efficiency of memory resources were responsible for developmental progress in reasoning (Bjorklund & Harnishfeger 1990). Of course, there are variables other than memory that change with age and may therefore be responsible for improvements in reasoning.

Another more fundamental problem with correlational evidence is that it does not directly address the issue of the necessity of memory for accurate reasoning on *specific* problems (Reyna & Brainerd 1990). Memory and reasoning tests might correlate simply because of individual differences; some types of children may tend to score high on both memory and reasoning tests, whereas other types may tend to score low. The two would then correlate, but not because memory is necessary for reasoning. Only analyses of memory-reasoning relationships at the level of individual problems are relevant to this question. As we have already seen, such analyses do not support memory necessity.

Because of the weaknesses of correlational designs, dual-task interference effects were advocated as providing more convincing evidence favoring the resources hypothesis. Typically, subjects in a dual-task experiment perform a motor task, such as finger tapping, while performing a reasoning task, such as mental arithmetic or transitive inference. Performance in both tasks usually declines, relative to a baseline condition in which no concurrent task is performed. The resource explanation for this decline is that, because memory capacity is limited, the capacity used by one task is no longer available for the other task. Developmental studies showed that the magnitude of the decline varied with age, being greatest in early childhood and old age, which was taken to reflect developmental differences in memory capacity (Bjorklund & Harnishfeger 1987; Guttentag 1989; Hasher & Zacks 1988).

Independent assessments of developmental differences in the memory-capacity variable, however, failed to reveal a consistent pattern across age (Dempster 1985; Reyna & Brainerd 1989b). This was especially true for the age range during which reasoning was undergoing rapid development. At the same time, the concept of memory capacity was being challenged in the adult literature, in some cases using the dual-task paradigm (e.g., Navon 1984). Finally, Brainerd and Reyna (1989) and Reyna and Brainerd (1989b) offered an alternative account of dual-task interference effects using the classical concept of output interference from the serial learning literature (Hull 1943).

According to this alternative account, outputting a response produces off-task noise, including response selection, scheduling, and feedback effects. For example, when subjects recall a list of previously studied words, a bottleneck occurs when those words are retrieved into consciousness more rapidly than they can be spoken. Subjects must then select and schedule the output of individual responses from among the retrieved items, and the auditory feedback produced by earlier words in the queue may interfere with the recall of later words (Brainerd, Reyna, Howe, & Kevershan 1991). Even when responses can be physically accomplished simultaneously, such as finger tapping and mental arithmetic, response selection, scheduling, and feedback from one task impinge on the other. The childhood game of rubbing one's stomach in a circular direction while rubbing one's head in the opposite direction is a familiar example. Although children have more difficulty with such tasks than adults, there are individual differences among adults in the ability to coordinate simultaneous outputs, and this ability declines in old age. Thus, output interference offers a coherent account of developmental data on dual-task interference, but, unlike capacity, it is also compatible with findings of memory independence.

THE NATURE OF REASONING

Other challenges to memory necessity and constructivism came from some reconceptualizations of reasoning itself. Reviews of contemporary theories of reasoning usually begin with the work of Piaget and his colleagues (e.g., Reyna & Brainerd 1990, in press). The influence of Piaget is undeniable; the paradigms he introduced and the phenomena that they elicited continue to be central to the study of reasoning. However, the explanations that Piaget offered for those phenomena have, for the most part, been supplanted by information-processing accounts (e.g., Bjorklund 1989; Siegler 1991).

Information-processing theories characterized reasoning as computation, bringing to bear analogies to machine intelligence. A rich assortment of concepts associated with intelligent machines became available, including the aforementioned dichotomy between short- and long-term memory. Theorists of cognitive development have also exploited ideas such as storage versus retrieval (Brainerd 1985b), processing speed (Kail 1993), knowledge structures and procedures (e.g., Bjorklund 1987), and so on. Most information-processing explanations of reasoning, however, revolve around the supposed capacity limitations of short-term (or working) memory (Baddeley 1986; Bjorklund 1989; Case 1985).

Inevitably, the memory-independence effect and related findings encouraged a closer examination of reasoning processes in a variety of tasks, including probability judgment, conservation, class-inclusion reasoning, transitive inference, and decision making. In the course of this work, it gradually became apparent that there was no convincing evidence that reasoners relied on the richly specified representations of problem information, as had been assumed by information-processing theorists. In numerical problems, for example, memory for simple qualitative distinctions—say, between more and less—were sufficient to predict reasoning performance (Reyna & Brainerd 1991a). Moreover, reasoners seemed to prefer more economical representations, especially when those representations afforded quick and accurate solutions (Brainerd & Reyna 1990b). At the same time, the literature in psycholinguistics supplied a steady stream of experiments, ranging from word-level metaphors to full-blown narratives, in which subjects seemed to encode, store, and remember the *gist* of verbal information, (e.g., Clark & Clark 1977; Kintsch 1974; Reyna 1981).

Because numerical tasks appear, on the surface, to be inextricably tied to the representation of precise quantities, such tasks provided some of the more striking examples of reasoning based on gist—i.e., qualitative representations. Piaget, for example, characterized probability judgment as involving quantitative compensation operations in which reasoners should process ratios of numerical frequencies (see also Surber & Haines 1987). So, for example, if subjects are asked to predict which of two containers provides the best chance of winning (i.e., of drawing a target), they should compute the ratio of targets to the total (or to nontargets). Piaget thought that the ability to make such computations was a by-product of a more general logical competence. For the information-processing theorists, however, competence was essentially quantitative (Reyna & Brainerd in press; Siegler 1981). Both traditions, however, stressed that, at the level of performance, mature probability judgment relied on the processing of numerical frequencies.

Despite the unanimity of theoretical opinion, the data suggested that probability judgments were based on vague distinctions such as which sample had more targets. Ratios, or numerical differences between targets and nontargets, were generally ignored (Brainerd 1981; Offenbach, Gruen, & Caskey 1984). Of course, such nonnumerical strategies have limitations. If samples sizes differ, estimating which sample has more targets can sometimes deliver erroneous responses (Reyna & Brainerd in press; Surber & Haines 1987). For example, if one sample has two targets and one nontarget and the other sample has three targets and six nontargets, the first sample has the better odds of producing a target on a random draw—although it has fewer targets. Although some studies showed that subjects adopted more elaborate ratio strategies when simpler strategies were inappropriate, it was generally found that these tasks were solved nonnumerically (Callahan 1989; Offenbach et al. 1984).

In another task involving numerical probabilities and outcomes, Reyna and Brainerd (1989a) showed that subjects reduced both probabilities and outcomes to qualitative categories. One group of subjects was presented with the well-known dread disease problem of Tversky and Kahneman (e.g., 1981). They were told that a disease that was expected to kill 600 people was about to strike the United States. They had a choice between two programs, one of which would save 200 people with certainty, whereas the other program offered a 1/3 chance of saving 600 people and a 2/3 chance of saving no one. This is the gain-frame version of the problem; outcomes are treated as positive increments with respect to the status quo. A different group received the loss-frame version: a choice between 400 people dying with certainty versus a 2/3 chance that 600 people would die and a 1/3 chance that nobody would die. Here, outcomes are treated as decrements with respect to the status quo. Note, however, that the gain and loss problems are equivalent. If 600 people are expected to die, then 200 saved leaves 400 dead, and so on. Despite the numerical equivalence of the gain and loss versions of the problem, the typical result is that people choose the certain option in the gain frame, but the risky option in the loss frame. This flipflop is called a framing effect.

Theories of the framing effect, and related anomalies of choice, generally posit a psychophysical function for quantities—probabilities and outcome values—through which objective quantities are transformed into subjective quantities (e.g., Tverksy & Kahneman 1986). The psychophysical functions introduce systematic deviations from objective values, much as perception of the intensity of light or sound is subject to predictable distortions. The observed anomalies of choice are accounted for by these distortions in the perception of quantities. The framing effect, for example, could be predicted by nonlinearities in the function that transforms probabilities (Kahneman & Tversky 1979). Although contemporary theories dispute the nature of the functions that transform objective quantities, they generally share the assumption that choices are determined by quantitative processing (Hogarth 1980).

Reyna and Brainerd (1989a) proposed that the data could be explained by assuming that processing was nonquantitative. Their specific proposal was that reasoners encode representations at varying levels of precision, that those representations can be ordered with respect to precision, and, most important, reasoning gravitates to the lowest, least precise, level in this "hierarchy of gist." Thus, in this view, quantitative information could be characterized in multiple ways that roughly correspond to scales of measurement. Beginning at the top of the hierarchy, ratio representations preserve numerical information with exactitude. Several steps below ratio scales, ordinal representations capture relative magnitude, but fail to preserve numerical differences (e.g., rank ordering probabilities according to which is largest, next largest, and so on). At the lowest level of representation, nominal or categorical, only the presence or absence of quantity is represented, for example, whether some lives are saved or no lives are saved (Reyna 1990a).

Reyna and Brainerd (1989a) claimed that, although problem information was represented at multiple levels, reasoning operated at the lowest level that would allow one to accomplish the assigned task, such as determining a preference between options or estimating the worth in dollars of lottery tickets with various odds. Obviously, response constraints in the latter kind of task rule out the processing of purely categorical representations; the response must be in dollars, rather than being a binary choice. Consequently, Reyna and Brainerd incorporated "task calibration" as a constraint on the level of representation (as well as on the functional advantages of relying on gist). The application of these elementary principles was straightforward, allowing Reyna and Brainerd to account for framing effects and to predict new phenomena.

For framing problems, the nature of the task provides minimal constraints, requiring only binary choice. Therefore, choices should be based on representations at the lowest possible level—the categorical level—in the hierarchy of gist. If quantities on both dimensions (probabilities and outcomes) were represented either as "some" or "none," the problems would boil down to something like choosing between, in the gain frame, saving some people versus saving some people or saving none; and, in the loss frame, some people dying versus some people dying or nobody dying. Probability is represented categorically because options are either *certain*, because there is only one possible outcome, or *uncertain*, because there is more than one possible outcome. Such representations were taken to be the gist, or the "bottom line," of the problem on which decisions turned. This was an extension of the term "gist" beyond its usual psycholinguistic meaning to numerical information.

It is easy to appreciate how subjects with such gist representations would prefer the option of saving some people (as opposed to none) in the gain frame, but shift to the option in which nobody died (as opposed to some dying) in the loss frame. Hence, standard framing effects could be explained. Crucially, these effects were obtained in adults, suggesting that mature reasoners were not solving the problems in the logico-quantitative manner posited by most theories of cognitive development. Similar results were subsequently obtained with adults in other reasoning paradigms (see Reyna & Brainerd 1991a, 1993).

Reyna and Brainerd (1989a) also produced experimental evidence challenging the traditional claim that the presence of numbers in the problem information is essential to these effects. They simply removed all of the numbers from the standard problems, replacing them with vague phrases such as "some people." Framing effects were not only obtained, but they were larger in magnitude when the numbers were absent than when they were present (see also Reyna 1990b). Thus, numerical information was not only unnecessary for these effects, but it tended to mask them rather than amplify them. More generally, this nonnumerical framing effect demonstrated that crude qualitative distinctions were sufficient to produce the standard pattern of preferences.

SUMMARY

Since the mid-1980s, research on the relationship between memory and reasoning, as well as research on reasoning per se, has converged on the conclusion that detailed nuances of problem information are not central to reasoning. Numbers, for example, were found to be unnecessary in the sense that memory for numbers was unrelated to reasoning and in the sense that numbers could be removed entirely from standard problems without eliminating standard effects. Analyses of reasoning revealed that qualitative distinctions—more versus less in probability judgment, some versus none in framing problems—accounted for well-known empirical patterns. Moreover, reasoners tended to operate on representations that were at the lowest level of precision that permitted a task-relevant response. These findings contradicted predictions of the dominant theories of the day, Piagetian and information-processing theories most especially. At the same time, findings that had seemed supportive of these dominant theories were shown to have alternative interpretations.

THE INITIAL MODEL

Theoretically speaking, the picture that emerges from the preceding section is an ambivalent one. On the one hand, the fact that findings such as memory independence and nonnumerical "numerical reasoning" are so counterintuitive makes them intrinsically interesting. But, counterintuitiveness is troublesome. Findings that are at odds with established ideas incur a heavy burden of explanation. Fuzzy-trace theory is the product of our efforts to provide an explanation for these counterintuitive findings. In the present section, we summarize how its core assumptions evolved.

The first attempts to bring theoretical order to these findings were contained in two papers that appeared in 1990 (Brainerd & Reyna 1990b; Reyna & Brainerd 1990). The crux of our approach consisted of seven related principles: gist extraction, fuzzy-to-verbatim continua, the fuzzy processing preference (intuition), reconstruction in short-term memory, output interference, resource freedom, and ontogenesis. The principles were designed to take account of then-recent research which, as we have discussed, challenged traditional views. However, we also drew on classical findings that had provided the original foundation for traditional views, including learning sets (Harlow 1949), discrimination transfer (Kendler & Kendler 1962), sentence verification (Reder 1982), memory for discourse (Kintsch 1974), judgmental heuristics and biases (Kahneman, Slovic, & Tversky 1982), memory suggestibility (Ceci, Ross, & Toglia 1987; Loftus 1979), interference (Hasher & Zacks 1988), and logical reasoning (Falmagne 1975; Trabasso 1977). We now consider each principle in turn.

Gist extraction is the notion that as children or adults encode informational inputs, they extract senses, patterns, and meanings, from them. This principle was supported by research indicating that subjects encoded global patterns or relationships on many reasoning tasks (Reyna 1981, 1985). This extracting of the gist of information had been demonstrated in transitive inference (Brainerd & Kingma 1984), the sense that objects got smaller and smaller as they were presented; in probability judgment, that one container had more targets than the other (Brainerd 1981); in class inclusion, that one class was larger than the other (Brainerd & Reyna 1990c); in framing, that one option involved saving some lives and the other involved possibly saving no one (Reyna 1990a, 1990b; Reyna & Brainerd 1989a); and so on. Research also indicated that subjects encoded more exact representations that, although they were not used much in reasoning, were retrieved to answer memory questions. Thus, a range of representations varying in precision, *fuzzy-to-verbatim continua*, were available for reasoning.

Despite the availability of multiple representations, in paradigms ranging from transitive inference (Reyna & Brainerd 1990) to decision making (Reyna 1990a), reasoning tended to operate on the less exact representations of gist, suggesting a *fuzzy-processing preference*. Mental arithmetic also illustrated *reconstruction in short-term memory*; subjects would rather reconstruct addends from their solutions to problems than retrieve them from memory (Brainerd & Reyna 1988a). The notion that memory limitations were central to explaining reasoning was further challenged by reinterpretations of dual-task effects in terms of *output interference* (Brainerd & Reyna 1989; Reyna & Brainerd 1989b). The reasoning-remembering independence effect, replicated in all of the major reasoning paradigms in cognitive development, favored the principle of *resource freedom*, that reasoning did not draw on a limited pool of generic resources (Brainerd & Kingma 1984, 1985). Finally, these principles identified dimensions of *ontogenesis*. Subjects of different ages varied in the representations and processes used in reasoning, as well as in their reliance on reconstruction and sensitivity to output interference (Brainerd & Reyna 1990b; Reyna & Brainerd 1990). Notably, older children and adults were prone to reason intuitively, contrary to virtually all theories of cognitive development.

As these seven principles make clear, although fuzzy-trace theory profited from earlier theoretical advances (some of which were incorporated into the theory), we chose to deal with empirical anomalies by introducing new theoretical concepts. We recognize, however, that it is sometimes possible to handle such anomalies by broadening traditional theoretical concepts (e.g., Bjorklund & Harnishfeger 1990; Salthouse 1990). We will briefly indicate why we felt compelled to break with tradition, and, once we accepted these new ideas, how their integration led to a different metatheoretical stance.

Our foremost consideration was that the theory-data conflicts were too basic. We might have attempted to retain the idea of limited capacity, for example, by assuming that there were specified pools of resources, each of which had limited capacity, as opposed to generic resources. Short-term memory could have been broadened to include reconstruction and other reasoning operations; and so on. As might be expected, broadening theoretical concepts in these ways reduced both predictive power and conceptual clarity; surprisingly, however, it also produced additional contradictions. For example, it seemed untenable to assume that memory limitations were the driving force behind higher cognition (Simon 1975) and, yet, accept that direct tests revealed no empirical relationship between memory limitations and higher cognition. Moreover, limited capacity, and other distinctions between short- and long-term memory (e.g., reconstruction) were core features of the "modal model" of information-processing (Schwartz & Reisberg 1991), just as logico-quantitative reasoning operations were core features of dominant theories of cognitive development (Siegler 1991). It became apparent that giving up core features of traditional theories amounted to giving up the theories. Thus, we began a search for alternatives to both the Piagetian and information-processing perspectives.

We found the alternative that we were looking for, one that was consistent with the empirically motivated principles discussed earlier, in the field of mathematics. The framework that we adopted—intuitionism—was originally developed in the context of the foundations of mathematics. Early in this century, three schools of thought arose regarding the foundations of mathematics. In content and chronology, they roughly correspond to three theoretical approaches to cognitive development. The first, *logicism*, assumed that mathematics reduces to "doing logic." Russell (1903), for example, attempted to prove that all the theorems of mathematics could be derived by defining mathematical concepts (e.g., numbers) in logical terms, and then applying logical rules to these definitions. Similarly, for Piaget (e.g., 1949), reasoning consisted of applying logical rules to "premises" (i.e., problem information).

The second school, *formalism*, was promulgated by Hilbert (e.g., 1923). Mathematics, in this view, consisted of manipulating symbols in accordance with certain rules, often compared to playing chess. It is not surprising that formalism bears some resemblance to information-processing theories of cognition inasmuch as it influenced early information theory (Broadbent 1958). Models of chess playing are prototypical of information-processing approaches to cognition, although most models are not directly implemented as computer programs (Schwartz & Reisberg 1991).

The third approach, *intuitionism*, was developed by Brouwer (1952) and Heyting (1959). Brouwer and Heyting criticized the rigidity of logicism and formalism—in particular, their inability to capture the fluidity and imaginativeness of mathematical thought. They defined an intuition as consisting of generic information (a basal entity) combined with an invention rule, from which new information could be generated. The goal for intuitionists, one that we share, was to precisely characterize an inherently imprecise process, the doing of mathematics.

In its beginnings, fuzzy-trace theory was a sort of psychological implementation of mathematical intuitionism. We interpreted the generic entities manipulated in thought as gists, and the processes of thinking as fluid and flexible. In contrast to logic and computation, we sought to capture the "rigorous sloppiness" of human cognition (Argyris 1988). The ability to tolerate noise and imprecision, for example, has long been recognized as a feature differentiating human versus machine intelligence (Estes 1980; von Neumann & Morgenstern 1947). Reasoning, therefore, is not, in its essentials, the rigid application of logical rules, nor the processing of exact data. Instead, it mostly consists of the flexible application of generic principles to imprecise representations.

If the natural habit of mind is to think intuitively in this sense, it followed that capacity limitations on the short-term recall of exact information would seldom impinge on reasoning. The tendency toward intuitive reasoning obviated the need to rely on verbatim memory for the exact form of inputs. The availability of multiple representations conferred flexibility on reasoning; a variety of representations (and of operations to process those representations) could be used to solve reasoning problems. As a construct differentiating short- from long-term memory, capacity was unnecessary. Many of the phenomena that had been explained by capacity could be accounted for by output interference or by the distinction between gist and verbatim representations. For example, verbatim memory becomes inaccessible more quickly than memory for gist, a fact well known in psycholinguistics (Clark & Clark 1977; Kintsch 1974). Therefore, the transience of certain kinds of memories (verbatim versus gist), as opposed to the nature of the memory store (short- versus long-term), was sufficient to provide a coherent account of differences between immediate and delayed memory tests. Since these constructs had all been necessary for other purposes, and their application allowed other constructs to be deleted, parsimony was increased despite greater coverage of disparate phenomena.

Further, the functional advantages associated with processing gist explained why mature reasoners could prefer intuitive reasoning (Brainerd & Reyna 1990b; Reyna & Brainerd 1990). These included the memorability of gist, the ease of processing it, the ability to extract gist patterns as information is encoded, and the ability to capitalize on tradeoffs between effort and precision without necessarily sacrificing accuracy. Neither was accuracy achieved by spurious means. Intuitive reasoning met criteria for logical competence (Reyna & Brainerd 1990). As Reyna and Brainerd (1991a) concluded, "it makes sense that cognition would be engineered around representations that are simple, flexible, efficiently processed, and that tend to be accessible in memory" (p. 49).

The identification of intuitive or "fuzzy" processing with *accuracy* was in contrast to traditional notions of accuracy in logical and computational theories. Indeed, we were often advised to eliminate the term "fuzzy" because it connoted inaccuracy. Of course, that connotation was the result of the implicit acceptance of logical and computational metaphors for reasoning. We retained the term "fuzzy" to underscore the fundamental differences between intuitionism and its alternatives.

CONSOLIDATION

BEYOND INDEPENDENCE

The initial model was followed by a consolidation period. The empirical base for fuzzy-trace theory became broader and more differentiated. The memory-independence effect was replicated and extended (Brainerd & Reyna 1992a), measurement issues were addressed (Brainerd & Reyna 1992b; Howe, Rabinowitz, & Grant 1993), and other laboratories obtained similar findings (Chapman & Lindenberger 1992; Fisher & Chandler 1991; Rabinowitz, Howe, & Lawrence 1989). By 1992, we were able to distinguish two sources of independence: (a) independence because reasoning performance is based on gist representations, but memory performance is based on verbatim representations and (b) independence because both reasoning and memory performance are based on gist representations, but retrieval errors are random, and therefore independent of one another (Reyna 1992a).

In general, we found that independence between memory and reasoning could be predicted by the specificity of the questions that were asked. In accordance with the fuzzy-processing preference, gist representations were used to answer questions except when verbatim specificity was required (providing verbatim representations were still accessible). As it happens, memory questions typically demand verbatim specificity, whereas reasoning can typically be accomplished with gist. In transitive inference, however, memory and inference questions are simply directional (e.g., Is the red stick longer than the white stick?) and both can be answered by remembering linear gist (the direction of "flow") and determining the order of items in the array. In fact, each combination of representations for reasoning and remembering leads to a clear prediction about their relationship, although some combinations are more likely than others (due to the fuzzy-processing preference; see Reyna, 1992a, Table 3.1). Thus, Reyna (1992a) concluded that reasoning-remembering independence was not intrinsic, but depended on the representations tapped in each task.

In addition to question specificity, Reyna (1992a) identified other factors that affected which representation, verbatim or gist, was used in a task. Predictions were based on whether representations of either type were retrieved (read out) versus reconstructed (Table 3.2). Thus, memory performance could be the product of, at one extreme, reading out verbatim traces, or at the other extreme, of reconstructing from gist. Although fuzzy-trace theory incorporated reconstructive processes in memory performance, we did not claim that memory was essentially constructive, as had Piaget (e.g., Piaget & Inhelder 1973) and others (e.g., Paris & Lindauer 1977).

Specifically, Reyna (1992a) pointed out that the findings of independence contradicted predictions of constructivism because tests of dependency are not directional. That is, positive dependency should be detected whether memory affects reasoning (necessity) or reasoning affects memory (constructivism). Instead, independence predominated, suggesting that memory was not essentially constructive. However, the theory stipulated specific conditions that ought to elicit reconstruction and, hence, ought to result in positive dependency. Preliminary results in that direction were reported from research on attitude change in which memories for pro and con arguments, after a delay, were associated with changes in attitudes (Reyna 1992a).

Also in 1992, Reyna and Brainerd addressed the issue of whether gist representations are extracted from verbatim representations or whether they are stored in parallel. Earlier, Brainerd and Reyna (1990b) had discussed two contrasting positions. According to one view, gist is derived from verbatim traces via inferential or interpretative processes (e.g., Glucksberg & Danks 1975). The other view is that gist is apprehended separately from verbatim traces in a process more like recognition than extraction. In other words, reasoners would have a variety of patterns stored in long-term memory whose instantiations would be recognized in problem information. It is well known, for example, that young children spontaneously encode relative numerosity when presented with objects that vary in number (e.g., Brainerd & Gordon 1994). In 1990, however, available data did not decisively favor one of these positions over the other.

Subsequently, however, it became clear that gist representations could be stored without passing through a preliminary stage of verbatim representations. Encoded representations of inputs (e.g., Farmer Brown owns 12 cows, 7 sheep, and 3 horses) could be used simultaneously to store verbatim traces (e.g., "12 cows"), and to access gist (e.g., "most") and interpret it (e.g., "cows are most"). Empirically, the parallelism scenario was supported by a number of findings indicating that verbatim and gist traces of the *same inputs* were functionally dissociated in memory (Brainerd & Gordon 1994; Reyna & Brainerd 1992; Reyna & Kiernan 1994b). In addition, the serial verbatim-first model—that verbatim representations are processed for gist and then expunged—had been called into question (e.g., Alba & Hasher 1983).

As more tasks were analyzed, it also became apparent that parallelism was supported by the fact that simple patterns or gists (e.g., "most") recurred across problems (Reyna & Brainerd 1991a). These patterns were available in long-term memory, and could be recognized before all the verbatim information was encoded, just as a familiar face can be recognized that is half in shadows. For example, in transitive inference, children picked up on the pattern that objects got smaller and smaller before all of the premises (e.g., The red stick is longer than the white stick; the white stick is longer than the blue stick; and so on) were presented (Brainerd & Reyna 1992b). Thus, gist could be recognized in parallel with the encoding of verbatim traces because certain patterns (e.g., "most," "smaller and smaller") were already available in long-term memory, and aspects of the problem (e.g., that objects varied in number or decreased in size) merely cued their retrieval (Reyna & Brainerd 1991a).

We continued to explore the notion that dissociated verbatim and gist representations were stored in parallel (Brainerd & Reyna 1993b). We thought that such dissociation could be detected in three ways. First, it could be detected by assessing the relationship between reasoning and remembering within a task (*horizontal independence*). Second, it could also be detected by assessing the relationship between age differences in memory and age differences in reasoning (*vertical independence*). Third, a factor's effect on memory could be compared to its effect on reasoning (*experimental independence*). Our initial research had focused on horizontal and vertical independence. However, the idea that memory and reasoning tapped dissociated representations implied that two types of experimental independence should also be found. Single dissociations, factors that affect one of the two measures (either of memory or reasoning), but not the other, and double dissociations, factors that affect memory and reasoning in opposite ways, were subsequently identified for several tasks (Brainerd & Gordon 1994; Brainerd & Reyna 1993a, 1993b, in press; Reyna 1995; Reyna & Brainerd 1993, in press; Reyna & Kiernan 1994b).

For example, certain manipulations (e.g., external stores of problem information) were identified that improved memory for problem information without affecting reasoning, and other manipulations were identified that had the opposite effect (e.g., removing misleading visual illusions at test improved reasoning but left memory unaffected). Most interestingly, under certain circumstances, factors that enhanced memory for problem facts actually depressed reasoning performance. Increasing the accessibility of numerical information led subjects down the garden path in class inclusion, for instance (e.g., Farmer Brown has 10 animals, 7 cows and 3 horses. Does he have more cows or more animals?). Correct reasoning was based on the ability to ignore numerical information (10, 7, 3) in favor of the qualitative gist that the classes form an inclusion hierarchy (everything is included in the class of animals) (Reyna 1991; Reyna & Brainerd 1993). In causal reasoning, when memory for premises was improved, children incorrectly rejected inferences because they could clearly remember that the inferences had never been presented (Brainerd & Reyna 1993b). Similarly, children remembered premises expressing spatial relationships more accurately than premises expressing linear relationships, but they performed worse on spatial inferences than linear inferences (Reyna & Kiernan 1994b).

With these findings of interference, we had come full circle with respect to the necessity hypothesis. The benefits of better memory for reasoning had given way to potential costs. Had we continued to accept the traditional view that memory and reasoning worked hand-in-glove, we would not have considered the possibility that memory and reasoning performance could be driven in opposite directions. Once memory and reasoning were dissevered, however, and different representations were posited for each task, it became obvious that these representations could react differently to the same manipulations.

REASONING REVISITED

Following its initial formulation, fuzzy-trace theory was also applied to a greater number of paradigms and, as suggested above, more detailed models of reasoning were proposed (e.g., Brainerd & Reyna 1990c, 1993b; Reyna & Brainerd 1991a, 1991b, 1993). The gist representations used in a variety of tasks were specified, along with the reasoning principles that operated on those representations (e.g., Reyna & Brainerd 1991a, see Tables 2 and 3). These representations and principles were broadly relevant across problems that differed in superficial detail. For example, the gist of relative magnitude ("one has more") is used in probability judgment, moral reasoning, decision making, and class-inclusion reasoning. Based on commonalities across various reasoning tasks, we were able to identify four sources of cognitive interference: interference from inappropriate verbatim or gist representations, from irrelevant reasoning principles, and from processing complexities.

The ability to recognize underlying gist, and to inhibit interference from superficial verbatim details, was found to be a factor in individual and developmental differences in reasoning (e.g., Dempster 1992; Cooney & Swanson 1990; Reyna & Brainerd 1991a; Schmidt & Welch 1989). Specifically, in tasks ranging from face recognition to causal inference, younger children were more likely than older children and adults to rely on superficial details, as opposed to global patterns (e.g., Brainerd & Reyna 1993b; Carey & Diamond 1977; Liben & Posnansky 1977; Perner & Mansbridge 1983). In discrimination learning, for example, children can be trained to pick the "middle" stimulus, say a circle measuring 7" in diameter, as opposed to a 5" or 9" circle. Given a transfer set consisting of a 7", a 9", and an 11" circle, adults and older children shift to the 9" circle, but younger children persist in choosing the verbatim training stimulus (Reyna & Brainerd 1992). This change in the representational basis for reasoning—the "verbatim-to-gist shift"—seemed to occur between early and middle childhood (Brainerd & Reyna 1993a; Reyna 1992a).

We differentiated the tendency to rely on verbatim versus gist representations from the ability to store gist representations, which was present early in life (e.g., Wagner, Winner, Cicchetti, & Gardner 1981). We also realized that, despite young children's tendency to rely on verbatim details, their ability to do so is limited in specific ways. For example, children who cannot subtract cannot base their reasoning on precise numerical differences (cf. Moore, Dixon, & Haines 1991). As children become more adept at processing precise differences, however, they are less likely to apply them in reasoning. Thus, the verbatim-to-gist shift implies that, when both verbatim and gist representations are available, younger children are more likely to rely on the former (Brainerd & Reyna 1993b; Reyna 1992a). During this same period, overall accuracy improves, consistent with the functional advantages of gist noted earlier.

The shift to gist-based reasoning, however, does not guarantee accuracy. Reasoners must also inhibit another source of interference: competing gist representations (Reyna 1991; Reyna & Brainerd 1991a, 1991b). In many cases, these competing representations act as retrieval cues for inappropriate reasoning principles. Reasoners who have difficulty inhibiting processing of these inappropriate representations and principles are led into systematic reasoning errors. In class-inclusion reasoning, for example, the gist of relative numerosity (e.g., "more cows than horses") prompts retrieval of the numerosity principle (i.e., the set with a bigger number has more) as opposed to the cardinality principle (superordinate sets have more than proper subsets). Therefore, when asked the class-inclusion question "Are there more cows or more animals?," a question that seems to be about numerosity, children respond erroneously, i.e. that there are more cows. The ability to inhibit interference from inappropriate gist representations and from inappropriate reasoning principles was also found to vary developmentally (Reyna 1991; Reyna & Brainerd 1991a; Brainerd & Reyna 1990b, 1993b).

Reasoners who manage to circumvent interference from inappropriate representations and principles can still become bogged down in the mechanics of processing. For example, relationships among overlapping classes can become sufficiently complex that reasoning accuracy is compromised (Brainerd & Reyna 1990c; Reyna 1991). Because processing complexity is discussed in detail in a later section, we will say only two things about it here. First, we distinguished processing interference from failures of reasoning competence. Processing difficulties can be addressed with superficial manipulations, such as diagramming the presented information in order to keep track of elements of the problem. Reasoners who lack reasoning competence do not benefit from such manipulations. Also, processing complexity should not be confused with processing load. In class-inclusion reasoning, for example, "load" is increased by increasing the number of nonoverlapping classes that must be processed, but, because these classes are distinctive, this decreases processing complexity (Reyna & Brainerd 1991a; Brainerd & Reyna in press a).

Thus, Reyna and Brainerd (1993) concluded that successful performance involves "selecting from among many relationships given as background facts, retrieving some among many principles that could be applied to such relationships, and, finally, applying the principles coherently" (p. 105). Interference could occur at any of these steps. Reasoners might fix on the wrong representation, retrieve the wrong principle, or become bogged down in processing complexities. Paradoxically, errors occurred despite excellent memory for the problem facts and the ability to solve the problem correctly (e.g., Brainerd & Reyna 1990c; Rabinowitz et al. 1989; Reyna 1991). To be sure, we did not claim that memory failure or ignorance of relevant principles never contributed to reasoning errors. By teasing apart these sources of errors, however, we found it to be more likely that reasoners misperceived problems that they were capable of solving, or they failed to retrieve their knowledge, or they had difficulty implementing knowledge that they had retrieved.

Such results reinforced the intuitionist view that reasoning was a dynamic and variable process in which multiple representations are encoded in parallel, correct principles are cued with some probability, and processing is sometimes unreliable. This accounted for variability in children's reasoning (Reyna & Brainerd in press) and for lapses in adult reasoning (e.g., Reyna & Brainerd 1991b). By separating errors of representation, retrieval, and processing, we were able to pinpoint developmental differences in susceptibility to various forms of interference. Such analyses highlighted differences between fuzzy-trace theory and traditional theories. Contrary to Piagetian theory, for instance, we found evidence for early logical competence; unlike information-processing theories, precise representations and elaborate processes were seen as sources of errors; and, contrary to contemporary emphases on strategic processes, we stressed basic memory processes, such as retrieval and interference.

CURRENT ISSUES AND FUTURE DIRECTIONS

We, and others, have continued to apply fuzzy-trace theory to a broad range of phenomena. Rather than enumerate these applications, however, we will focus on recent developments in those selected domains in which theoretical principles have been most systematically investigated. These domains are suggestibility and false memories, judgment and decision making, the development of forgetting, and the development of retrieval. As will be seen, by pursuing highly specific hypotheses in separate domains, ironically, our ideas have become more theoretically integrated.

SUGGESTIBILITY AND FALSE MEMORIES

False Recognition. There has long been controversy about the accuracy of memory, the resolution of which has important implications for applied problems in eyewitness testimony, classroom instruction, and clinical psychology. As we have discussed, some theorists have assumed that memory is constructed out of our understanding of experience (Turvey 1974). Although the roots of constructivism in memory extend back to Piaget (e.g., Piaget & Inhelder 1973), Bartlett (1932), and beyond, the false recognition experiments of the 1970's seemed to provide an empirical foundation for earlier conjectures. The main result was that subjects often could not discriminate between presented sentences and unpresented sentences that preserved the gist of presented sentences; under some conditions, memory confidence was higher for the unpresented sentences (e.g., Bransford & Franks 1971). Therefore, it seemed that the accuracy of memory was in doubt because actual experiences of the world could not necessarily be discriminated from the world-as-understood.

If one accepted the data, they seemed to offer prima facie evidence for constructivism, including the implication that the accuracy of memory was suspect (Schwartz & Reisberg 1991). The data were not universally accepted, however. Most criticisms had to do with methodology (reviewed in Alba & Hasher 1983; see also Fletcher 1992). For example, researchers argued that superficial familiarity was confounded with gist consistency. Sentences that were true, whether presented or not, tended to contain words that had been presented, whereas false sentences contained novel words (e.g., Flagg 1976; Reitman & Bower 1973; Small & Butterworth 1981). These critics maintained that memory for surface features of sentences could explain misrecognition data.

Although there were disputes about the data, observers on both sides accepted the premise that, if the effects were real, memory had to be constructive. Fuzzy-trace theory introduced a new possibility, namely that the effects were real, but memory was not constructive. Reyna and Kiernan (1994b) modified instructions and stimuli to address the methodological criticisms of earlier work. The aim was to obtain unambiguous demonstrations of the classic effects in order to determine whether they were explained by constructivism. If subjects systematically misrecognized unpresented true sentences, constructivism would have entailed positive dependency between memory judgments for presented sentences and for unpresented sentences consistent with gist.

On the other hand, according to fuzzy-trace theory, two kinds of memory representations of actual experience are deposited—verbatim and gist—that are independent of one another (see also, Kreindler & Lumsden in press). Under conditions such as those typically used in sentence recognition experiments (e.g., immediate memory tests), verbatim memory remains accessible and governs recognition of presented sentences. However, systematic misrecognition of new semantically consistent sentences (e.g., true inferences) cannot be based on verbatim memory for the simple reason that these sentences were never presented. Subjects sometimes say "yes" to semantically consistent sentences, however, because they cue memory for the gist of what was presented (e.g., Ackerman 1992). Thus, under these conditions, recognition of old sentences and misrecognition of new semantically consistent sentences should be independent.

Reyna and Kiernan's (1994b) results confirmed the prediction that recognition of old sentences and misrecognition of new semantically consistent sentences were independent. These relationships could be manipulated experimentally by varying whether subjects were instructed to base judgments on gist as opposed to verbatim memory. Merely instructing subjects to base all judgments on gist converted the relationship between the *same* presented sentences and true inferences to positive dependency. Delaying the memory test until verbatim memory was no longer accessible had a similar effect. (Effects of delay on gist versus verbatim memory are well documented in the psycholinguistics literature, Clark & Clark 1977.) In addition, younger children, who were less likely to acquire an accurate verbatim representation, were also more likely than older children to display positive dependency.

Interestingly, both independence and dependency could be elicited within the same task. The verbatim memory test included true paraphrases, sentences that were synonymous with presented sentences but were worded differently, as well as true inferences. Like true inferences, true paraphrases were misrecognized more often than familiarly worded false sentences. Although judgments of presented sentences and true inferences were independent, judgments of true paraphrases and of true inferences were dependent. The latter were both consistent with the gist of presented sentences, but neither had been presented.

These results indicated that memory performance will sometimes conform to the expectations of constructivism, and sometimes will not so conform, depending on whether judgments are based on gist or verbatim memories. When subjects based judgments on their memory for gist, those judgments (i.e., of true paraphrases and true inferences) were positively dependent, as predicted by constructivism. For presented sentences and true inferences, this could be achieved simply by instructing subjects to base both judgments on gist.

However, if judgments of presented sentences and true inferences were based on separate memory representations, verbatim in one instance and gist in the other, those judgments were independent, which is not predicted by constructivism. Also contrary to constructivism, within the same task, different relationships were obtained for presented sentences and true paraphrases, despite the fact that they were synonymous. Each of these effects—of sentence type, instructions, delay, and age—could be explained in terms of whether judgments were based on gist versus verbatim representations.

Negative Dependency. The findings obtained by Reyna and Kiernan (1994b) were subsequently replicated with adults using similar materials (Lim 1993), and with children using different materials (Reyna & Kiernan 1994a). For instance, children misrecognized true interpretations of metaphors more often than false interpretations in a verbatim memory task (Reyna & Kiernan 1994a). Recognition of presented metaphors and true interpretations tended to be independent, but recognition of different types of interpretations tended to be positively dependent. Instructing children to base all judgments on gist produced positive dependency between presented metaphors and true interpretations. These findings with metaphorical sentences and the gists of those sentences, namely their interpretations, resemble Reyna and Kiernan's (1994b) findings for literal sentences.

Although the overall pattern of findings was similar across studies, differences in the degree to which verbatim memory was accessible were found. Lim's (1993) study with adults using literal sentences confirmed the developmental trend observed by Reyna and Kiernan (1994b). Verbatim memory was better in older subjects; adults were better able to discriminate presented sentences from gist, compared to children in Reyna and Kiernan (1994b). The metaphor studies further confirmed, as also observed in Reyna and Kiernan (1994b), that verbatim memory was better for certain stimuli. Consistent with previous claims about the memorability of metaphors, children were better able to discriminate presented metaphors from gist, compared to literal sentences. Together, these results indicate that verbatim memory would be maximized for adults judging metaphors and metaphorical interpretations. Thus, by using different age groups and materials (in combination with different instructions and delays), it was possible to manipulate the representations subjects used to make recognition judgments.

Thus, the Reyna and Kiernan (1994b) study led to a number of follow-up studies in which the representational bases for recognition judgments were varied. Reyna and Kiernan (1994b) had already found that instructions could induce subjects to base all judgments on gist, producing positive dependency. Although verbatim memory was more ephemeral than gist, theoretically, it ought to be possible to induce subjects to base all judgments on verbatim memory. If that were accomplished, recognition of old sentences and new semantically consistent sentences should be negatively dependent (for constraints on this prediction, see Brainerd & Reyna 1992b; Reyna 1992a). The key, however, was to find conditions in which verbatim memory could be readily accessed. According to the results presented above, adult memory for metaphors should provide such conditions.

Thus, a study using metaphors was conducted with adults that was virtually identical to the one conducted with children (Lim 1993; see also, Reyna in press).

As expected, the level of verbatim discrimination was excellent. In fact, unlike any of the earlier studies with either literal or metaphorical stimuli, there was no evidence of misrecognition of gist on immediate verbatim tests. Adults were apparently able to base their judgments solely on verbatim memory for metaphors, rejecting true and false interpretations equally (including those that were highly similar in surface form to the original metaphors). As noted, if recognition in both cases (for both presented metaphors and for interpretations) were based on verbatim memory, judgments should be negatively dependent. That is, subjects would both accept presented metaphors and reject interpretations based on memory for the exact wording of original sentences. Results generally conformed to this prediction.

Thus, we have moved beyond independence empirically and theoretically. Empirically, we have found that the degree to which subjects can discriminate presented stimuli from gist varies systematically with age, stimuli, delay, and instructions. Theoretically, the level of such discrimination indicates the accessibility of verbatim versus gist representations. This, in turn, can be related to the type of relationship—positive, independent, or negative—that is likely to be obtained between judgments of presented stimuli and of gist. In a verbatim memory task, when verbatim memory can be accessed regardless of the cue (i.e., for either presented or unpresented stimuli), negative dependency is obtained. If gist is accessed regardless of the cue, positive dependency is obtained. And, in that intermediate stage of verbatim accessibility, when the representation that is accessed changes with the cue (verbatim or gist), independence is obtained.

False Recognition Reversal. As we have discussed, false-recognition effects occur when subjects erroneously accept semantically consistent items as having been presented. There is wide agreement that such items are accepted because of their consistency with the gist of presented material. Fuzzy-trace theory differs from other accounts, however, in assuming that multiple dissociated representations, gist and verbatim, can be accessed in recognition judgments. If verbatim rather than gist representations were accessed for unpresented stimuli, as we have seen for metaphors, subjects should reject those stimuli rather than accept them. Although some stimuli spontaneously elicit verbatim representations more easily than others, we sought a paradigm in which we could gain experimental control over verbatim accessibility.

We decided to modify a lexical priming task in order to make it possible to prime verbatim representations (Brainerd, Reyna, & Kneer in press). In the usual priming task, subjects make lexical judgments (e.g., word-nonword judgments) about targets (e.g., "doctor") after having just been exposed to some other item. When that item is related to the target (e.g., "nurse"), the exposure primes the target because response time is decreased, compared to control conditions in which the item is unrelated to the target. Of course, in the standard procedure, there is no verbatim representation of the target word because judgments (e.g., whether the item is a real word or a nonsense syllable) are based on representations in semantic memory. Therefore, we asked subjects to *first* study a list of words, and then make old-new judgments on a memory test.

Although studying a list provided verbatim representations of targets, it was then necessary to prime those representations in order to manipulate verbatim accessibility. Reyna and Kiernan's (1994b) results indicated that presented items cued their verbatim representations. Therefore, we primed verbatim representations by presenting the targets themselves prior to related distractors (Brainerd, Reyna, & Kneer in press). At issue was the effect of verbatim priming on related distractors.

Many theories of priming would predict that studying targets should produce bigger priming effects and, hence, higher acceptance rates for related distractors. On the other hand, our predictions turn on the nature of the representations that are primed. Verbatim comparisons between related distractors and targets should produce wholesale rejection. This is because verbatim comparisons could only lead to two outcomes, a match or a mismatch, and related distractors were mismatches. Gist comparisons, in contrast, could produce acceptances for related distractors because of their global similarity to targets.

We found that, in the absence of verbatim priming, the usual false recognition effects were obtained. Judgments of related distractors tended to be based on gist, and acceptance rates were increased relative to unrelated distractors (as in Reyna & Kiernan 1994b). When verbatim representations were primed, however, related distractors were rejected. Naturally, related distractors were more effective cues for target representations than were unrelated distractors. This led to *higher* rejection rates for related than for unrelated distractors under verbatim priming conditions, called false-recognition reversal (Brainerd, Reyna, & Kneer in press): Once reminded of the verbatim target, subjects were better able to reject the distractor. Thus, changing the representation that subjects used for their judgments from gist to verbatim made it easier to reject related distractors than unrelated distractors, despite the greater similarity of related distractors to targets.

It should be noted that, consistent with what is known about the effects of delay on verbatim accessibility, lags of even a few words between targets and related distractors substantially weakened false-recognition reversal effects. After delays of over a week, false-recognition reversals were replaced by the usual false-recognition effects for the same stimuli. Also consistent with previous results regarding verbatim memory, older children showed bigger false-recognition reversal effects than younger children. Thus, the similarity of distractors to targets had opposite effects for different age groups, at different delays, and under different priming conditions, depending on the kind of memory representation that was used.

Interference and Suggestibility. In a recent paper, Reyna (1995) has argued that memory and reasoning phenomena can be organized in terms of a taxonomy of representations and tasks. This taxonomy provides an explanatory structure in which disparate phenomena can be related and, consequently, theoretical claims can be integrated. In this section, we briefly note the place of previously discussed phenomena in the taxonomy, and new applications to source confusions, suggestibility, and autosuggestibility.

In the taxonomy, there are two classes of questions that subjects might be asked and two classes of representations that subjects might use to answer those questions.

(Actually, we assume a continuum of questions and of representations, but will treat these as dichotomies for the sake of exposition.) Questions can vary in specificity, from requiring exact information (verbatim questions), such as the precise number of targets presented in a probability judgment task, to global judgments (gist questions), such as which sample had more targets. Verbatim or gist representations can be used to answer these questions. This implies that there are four possible types of errors: giving an inappropriate gist or verbatim response to a verbatim question, or giving an inappropriate gist or verbatim response to a gist question.

We have already discussed inappropriate responses to gist questions. For example, evidence for inappropriate verbatim responses in a gist task was obtained by Brainerd and Reyna (1993b), who found that children incorrectly rejected true inferences because the inferences had never been presented. That is, children interrogated verbatim memory, failed to find a match, and exited prematurely (as opposed to consulting gist memory). A more common error in reasoning tasks, however, is to rely on the wrong gist representation (e.g., in class-inclusion, probability judgment, and conservation). Consistent with the verbatim-to-gist shift in reasoning, errors of the first type—verbatim interference with reasoning—are generally confined to young children.

Just as responses can be limited to verbatim information, though the question calls for gist, gist can be misidentified as verbatim information, i.e., as having been presented. A number of familiar phenomena can be classified as gist interference with a verbatim task. These include standard false-recognition effects (Paris & Carter 1973), inferential remembering in free recall (Brown, Smiley, Day, Townsend, & Lawton 1977), say/mean confusions in which inferences are misattributed to text rather than to self (Beal 1990a, 1990b), and source confusions in which imagined actions are misremembered as having actually been performed (Foley, Johnson, & Raye 1983). Although these phenomena have each been explained by different theories, our analysis suggests that they can be grouped together under the same explanatory umbrella.

The constructivist explanation for verbatim memory errors is that experienced events are integrated with inferences and other elaborations that go beyond direct experience. In contrast, fuzzy-trace theory assumes that verbatim memories are not integrated with gist. Therefore, verbatim memory errors need not imply that subjects have forgotten what they have experienced. Instead, such errors can occur because gist is accessed, rather than verbatim memory (Reyna & Kiernan 1994b). Consistent with this interpretation, errors vary directly with the relative accessibility of gist versus verbatim representations, and factors that influence the accessibility of these representations, such as age, prior knowledge, and delay, have predictable effects (e.g., source errors increase after a delay or when clues about inferences precede memory questions, Ackerman 1992; Reyna in press a).

As children get older, they are more resistant to interference of all types, including gist interference with verbatim judgments (Dempster 1992). For example, although younger children are not subject to greater misrecognition of inferences under neutral cuing conditions (e.g., Reyna & Kiernan 1994b; Liben & Posnansky 1977), they are less able than older children to inhibit interference once inferences are cued

(Ackerman 1992; Beal 1990a, 1990b). Thus, because older children are better able to inhibit gist responses in verbatim tasks, and verbatim responses in gist tasks, there is *greater* dissociation between gist and verbatim representations with age, contrary to predictions by Piagetian theorists (e.g., Chapman & Lindenberger 1992; Prawatt & Cancelli 1976).

In addition to being examples of gist interference in verbatim memory tasks, standard false-recognition effects, inferential remembering, say/mean confusions, and certain source errors are all examples of "autosuggestion" (Reyna in press a). In autosuggestion, subjects confuse internally generated events with events that were actually experienced (Binet 1900; Ceci & Bruck 1993). Autosuggestibility has also been studied in the context of problem solving (Brainerd & Reyna in press a). When children are asked memory questions as they solve class-inclusion problems, they systematically confuse the number of items in the larger subset (e.g., cows) with the number of items in the superordinate set (e.g., animals). (In control conditions involving the same numbers, but nonoverlapping sets, children's memory is highly accurate.) The pattern of memory errors mirrors the pattern of reasoning errors: When asked the reasoning question about which class is bigger, children claim that the subset is larger than the superordinate set, and larger numbers are chosen for subsets than superordinate sets in memory tests. However, memory and reasoning were found to be independent, again because errors depended on retrieval, and different retrieval cues operated for memory versus reasoning questions. As in the other examples of autosuggestion, these effects in class inclusion can be classified as gist interference with verbatim memory (see Brainerd & Reyna in press a). (Note that children confuse superordinate sets with larger subsets, not larger subsets with smaller subsets, thereby disconfirming the usual linquistic ambiguity hypothesis.)

For both autosuggestion and suggestion, subjects misreport actual experience, but in suggestion the impetus is external. In a typical study on suggestion, subjects are presented with an event (e.g., slides of a car accident), and are later presented with misleading information about that event (e.g., that the car was blue rather than green). Reyna (1995) distinguishes suggestibility effects that have to do with memory for events themselves from effects that have to do with memory for source. Memory for events can be based on either gist or verbatim representations. Information about source, *as it is typically manipulated in experimentation* (e.g., Lindsay 1990; Lindsay & Johnson 1991), is likely to be represented only in verbatim memory. Thus, suggestibility effects can occur either because subjects fail to remember *what* was presented or because they remember what was presented, but not *where* it was presented, and erroneously attribute later information to the original experience. Unlike theories which treat all suggestibility effects as source confusions, this analysis implies that there are different kinds of effects that have to do with different kinds of representations.

Titcomb and Reyna (1995) note that four tenets of fuzzy-trace theory are particularly relevant to the interpretation of suggestibility effects. These tenets concern: (a) the nature of gist versus verbatim memory representations, (b) the independence of gist and verbatim memories, (c) the nature of forgetting (as the disintegration of traces), and (d) the differential forgetting rates for gist and verbatim memories. A number of researchers have used these tenets to explain suggestibility effects in children and adults (e.g., Brainerd in press a; Brainerd & Ornstein 1991; Cassel & Bjorklund 1994; Ceci & Bruck 1993; Poole & Lindsay in press; Poole & White 1993, in press; Reyna 1992a, 1995; Reyna & Kiernan 1994b; Warren & Lane in press). Based on suggestibility research, as well as the broader base of research on memory and its development, it is possible to identify conditions that should produce different types of suggestibility.

Specifically, misleading suggestions usually concern specific details, rather than gist (e.g., the color of a car involved in an accident; the kind of tool used in a break-in, etc.). If subjects have access to verbatim memories for original events, they will resist such suggestions. The specificity of verbatim representations, then, is key to blocking acceptance of contradictory details. If subjects rely on gist, on the other hand, misleading details do not necessarily contradict subjects' memories. The tendency to rely on gist may be increased by efforts to make sense of one's experience, for example, in order to assign guilt or innocence (e.g., Pennington & Hastie 1993). Therefore, subjects might accept misinformation because it is compatible with the substance of what they remember. Because gist and verbatim memories function independently, it is possible to accept misinformation—by relying on gist—despite accurate verbatim memories for original events (Reyna & Kiernan 1994b).

Which representation subjects rely on, verbatim or gist, depends on forgetting (e.g., Poole & White 1993; Warren & Lane in press). Forgetting rates are higher for verbatim (as opposed to gist) representations. Therefore, subjects are more likely to rely on verbatim representations immediately after original information is presented, but shift to gist after a delay (Reyna & Kiernan 1994b). Forgetting is not an all-or-none loss of a trace, however. Instead, because forgetting is the gradual disintegration of the features of a trace, pieces of a trace (e.g., source information) can become dissociated from one another (Reyna 1992a). Under certain circumstances, the pieces can also be redintegrated, or re-stored (e.g., Brainerd, et al. 1990). These general assumptions about the role of gist versus verbatim memory in studies of suggestibility imply certain specific predictions (Titcomb & Reyna 1995). First, direct blocking between original and misleading details should be observed under immediate conditions. That is, when original and misleading information are presented close together in time, and memory is tested immediately thereafter, verbatim memories for both events are still accessible. Their specificity forces an either-or choice between them. Chandler's (1991) data, for example, are consistent with this account; blocking was detected on an immediate test, but disappeared after a delay.

Under such conditions, in which verbatim memories directly compete, Loftus, Levidow, and Duensing's (1992) principle of discrepancy detection (which states that the poorer one's memory for presented material, the more susceptible one is to suggestion) would apply. The principle might also apply if the gist of both original and misleading information were to conflict—i.e., if the acceptance of one version of events entailed the rejection of the other version—but this has not been explicitly investigated (but see Ceci & Bruck 1993). The principle does not apply to verbatim and gist memories, however, because they typically do not directly contradict one another.

After a delay, original verbatim memories are gradually forgotten, and, so, they cannot compete with verbatim memory for misleading information (usually presented closer in time to the memory test). Thus, delays between original information and the memory test, and more recent presentation of misleading information, should favor observing suggestibility effects, and this, too, has been observed (e.g., Belli 1989; Cassel & Bjorklund 1994; Ceci & Bruck 1993; Titcomb & Reyna 1993; Warren & Lane in press). Consistent with this interpretation, factors that enhance verbatim memory for original events, such as early memory tests, decrease suggestibility after a delay (Brainerd in press; Warren & Lane in press).

Also after a delay, traces should become disintegrated so that it is possible to forget such aspects of an experience as its source, while remembering others. For example, witnesses might remember seeing an object in a certain location, but not when it was last seen there, or witnesses might know that certain events took place, but be unable to tell whether they remember the events or whether they were told about them. Such scenarios, in which witnesses can report the substance of events, but not where or when they learned about them, would be especially probable because source is typically a verbatim detail. Therefore, source should be forgotten more quickly than the substance of information, creating opportunities for suggestion (Reyna 1995). Evidence favoring this kind of "source confusion" explanation for suggestibility effects has been reported (e.g., Belli, Lindsay, Gales, & McCarthy 1994; Lindsay 1990; Lindsay & Johnson 1991).

Although traces disintegrate across forgetting intervals, they can be redintegrated, for example, by repeated questioning (Brainerd & Ornstein 1991). Warren and Lane (in press) have replicated this repeated-questioning effect using a suggestibility paradigm. After receiving the original and then the misleading information, subjects were given neutral questions about the original events. They showed significantly less suggestibility, compared to subjects who had received no questions after the original event. Similarly, Belli et al. (1994) found that suggestibility effects were not detectable on a final memory test after an interpolated test for original events (Experiment 1). Although the textbook view of forgetting is that performance declines over time, these findings indicate that decline is not inevitable. Through repeated questioning and other factors that spur redintegration, subjects can apparently recover previously forgotten memories.

In summary, susceptibility to suggestion depends on factors affecting the accessibility of verbatim memories for original events (e.g., Brainerd in press; Reyna 1992a). Hence, factors that increase verbatim forgetting increase suggestibility effects. Delaying misleading information increases its impact because competing original information is forgotten (e.g., Chandler in press). On the other hand, when subjects receive immediate tests for original events, which reduces forgetting of verbatim information, later suggestions have smaller effects (Warren & Lane in press). Repeated questioning after a delay also combats the effects of forgetting: Memory performance improves—and suggestibility effects decrease—on repeated tests (e.g., Brainerd in press; Poole & White 1993, in press).

Although gist representations of original events are accessible for longer periods of time, they lack specificity, and so they do not necessarily contradict misleading details. Therefore, despite robust memories for the gist of events, suggestions rejected at short delays might be accepted after long delays. As original events are forgotten, elements such as source can become dissociated from other aspects of the events. To the degree that source memories are verbatim, they will be forgotten more rapidly than the substance of events, producing confusions on later memory tests.

In short, suggestibility effects encompass a collection of phenomena, many of which have been investigated in basic research on memory (e.g., Brainerd in press; Brainerd et al. 1990). Factors that have been studied extensively in research on memory, such as delay, immediate testing, repeated testing, and so on, have analogous effects in research on suggestibility (e.g., Brainerd in press a). The theoretical interpretation of these effects explains both consistencies and apparent contradictions in the suggestibility literature. In particular, four principles of fuzzy-trace theory, that were derived from basic memory research, can be used to organize the variety of findings regarding suggestibility: that gist and verbatim representations of events function independently, that these representations vary in specificity and in forgetting rates, and that forgetting consists of the gradual fragmentation of memories (e.g., Reyna & Brainerd 1992).

JUDGMENT AND DECISION MAKING

As we have indicated, research in judgment and decision making has influenced fuzzy-trace theory from its initial formulation. Here, we review recent developments in three areas: (a) differences between judgment and decision making, (b) research on judgment, and (c) research on decision making. In the section on differences between judgment and decision making, we illustrate how such differences are predicted by fuzzy-trace theory's assumptions regarding task calibration. In the section on judgment, we explain how the complexities of inclusion relationships can produce a chain of errors in representation, retrieval, and processing, providing a unitary account of class-inclusion reasoning, conjunction fallacies, conditional probability judgment, deductive reasoning, and other tasks. In the section on decision making, we describe a series of experiments pitting predictions of traditional theories of decision making, which claim that numerical processing is central to observing decision phenomena, against those of fuzzy-trace theory. Specifically, we examine whether decision makers process numerical information, and whether such numerical information is either necessary or sufficient to produce standard decision phenomena. Finally, we discuss the implications of research in these three areas for conceptions of cognitive competence and rationality.

Task Differences. Judgments are distinguished from decisions primarily by the response that subjects must give, rather than the information that is presented. For example, given the same information about features of apartments for rent, such as size and location, subjects might be asked either to choose the best apartment (decision making) or to indicate the rent that they would be willing to pay for each apartment (judgment). The common-sense assumption is that judgments and choices should be closely related. For example, subjects should be willing to pay more for preferred apartments. Although judgments and choices are not entirely unrelated, inconsistencies between them have been amply documented (e.g., Hogarth 1980; Payne, Bettman, & Johnson 1992; Tversky 1969). Therefore, we will first discuss overall differences between such tasks, followed by more detailed analyses of specific judgment and decision-making paradigms.

Researchers have proposed different models for judgment as opposed to decision making, and there is general agreement that processing can be simpler in decision making (e.g., Wedell & Bockenholt in press; Fischer & Hawkins 1993; Payne, Bettman, & Johnson 1988; Tversky 1972; Tversky, Sattath, & Slovic 1988). Simplicity is usually evaluated with respect to ideal computational models, often drawn from statistics and probability theory. The question has been whether subjects process all of the relevant information in a sufficiently precise and elaborate manner, in accordance with quantitative rules. Although both judgment and decision making fall short of quantitative ideals, direct comparisons between these tasks indicate that decision making is more likely to elicit qualitative processing.

As Fischer and Hawkins (1993) have pointed out, traditional models do not fully account for task differences. Although there has been substantial progress is pinning down the differences in processing across tasks, the reasons for those differences are not clear. Since Simon (1956), researchers have attributed simplified processing to the memorial limitations assumed in the computer model of mind (e.g., Nisbett & Ross 1980; Payne, Bettman, & Johnson 1992). However, general assumptions about memory limitations do not explain why choice differs from judgment. Moreover, as in the problem-solving literature, decision makers engage in simplified processing when memory demands are few, that is, when a single decision is presented, there is little information to be processed, the information is written down, there is no time limit, and the relevant quantitative operations are easy to perform (e.g., dividing 600 by 3).

From the perspective of fuzzy-trace theory, differences between judgment and decision making are predictable because responses differ in precision (e.g., Reyna & Brainerd 1991b, 1992, in press). As we have discussed, the preference for fuzzy processing is constrained by the specificity of the response that is required in a given task. Thus, choices can be based on the simplest categorical representations, but these are insufficient to provide numerical responses, such as prices. This principle of task calibration also allows us to extrapolate beyond choice and pricing to other responses that vary in precision, such as rankings and ratings. Processing differences for such tasks, consistent with predictions of fuzzy-trace theory, have also been reported (e.g., Fischer & Hawkins 1993).

The level of representation that is used in judgment and decision making, however, does not depend solely on the form of responses. In a forced-choice situation, task considerations include the ability to discriminate between options, and, therefore, the nature of those options can influence the level of representation. (More general considerations include the ambiguity of quantitative information, and the functional consequences of assimilating nonidentical quantities; see Reyna & Brainerd 1992). For example, consider a choice between a 90% chance of winning $10 and a 10% chance of winning $100. Unlike the framing problems presented earlier, both options involve risk. Therefore, categorical distinctions between certainty and uncertainty are not possible. In order to differentiate between these options, higher-order distinctions (e.g., between more versus less) must be invoked. Thus, the nature of the options themselves affects the level of representation that subjects can use to generate a response.

Wedell and Bockenholt (in press) have applied fuzzy-trace theory to explaining inconsistencies between tasks, as well as to inconsistencies within a given task (e.g., choice) as the options change. In particular, Wedell and Bockenholt examined changes in decision anomalies when subjects are instructed to anticipate repeated plays. They noted that this manipulation reduces the difference between choice and pricing, as well as other decision anomalies, such as the Allais paradox (Keren & Wagenaar 1987; Wedell & Bockenholt 1990, 1992). Consistent with fuzzy-trace theory, Wedell and Bockenholt (in press) found that anticipation of repeated plays produces a shift from qualitative processing to quantitative processing, which reduced anomalies. The explanation for this shift is straightforward.

Suppose that subjects must decide whether to accept a wager offering a 50% chance to win $100 and a 50% chance to lose $50. In a single play, subjects will either win or lose, and they generally avoid the possibility of losing by refusing to accept the wager (e.g., Kahneman & Tversky 1979). Across many plays, however, outcomes vary, and the proportion of wins and losses will approach the stated probabilities. For multiple plays, therefore, subjects do not face the prospect of either winning or losing, but, rather, an aggregation across these outcomes. Thus, because categorical distinctions no longer apply for multiple plays, subjects make finer distinctions.

The difference between one-time versus repeated plays is illustrated by the biblical story of Solomon and the baby claimed by two mothers. In the biblical story, either one mother will receive the baby or the other will. Fifty percent of the baby, although it satisfies the expectation principle, is not a real possibility. Similarly, in a one-time choice such as the wager above, the decision maker will not receive a percentage of an outcome. Indeed, the categorical nature of outcomes is often what makes choices so difficult (e.g., Hogarth 1980). Such considerations as the form of response (e.g., choices versus numerical estimates), the nature of the options, the amount of uncertainty surrounding quantities, and functional criteria concerning the consequences of assimilating nonidentical amounts apply in designating the level of representation used in a task (Reyna & Brainerd 1991b, 1992). Thus, subjects operate at different levels of precision in different tasks.

Judgment. Our discussion of differences between judgment and decision making focused on differences in representations. As our earlier discussion indicated, however, performance is also affected by retrieval and processing. Of course, as we have suggested, representation, retrieval, and processing are interconnected. If subjects fail to encode relevant relationships, or irrelevant relationships are more salient than relevant ones, the correct reasoning principle may not be retrieved (Reyna & Brainerd 1991a). If the correct principle is not retrieved, manipulations that improve processing will not be effective (Brainerd & Reyna 1990c; Reyna 1991). The key to preventing this chain of cognitive mishaps is knowing what the question is. Although it seems a trivial matter to know what the question is, many judgmental errors fall into this category (e.g., Kahneman, Slovic, & Tversky 1982; Nisbett & Ross 1980). Subjects fix on a salient representation that seems to answer the question, except that it does not. In class—inclusion reasoning, we saw that representations of the relative magnitude of subsets (e.g., cows are more numerous than horses) are used to answer a question about inclusion relationships. In probability judgment, representations of the relative magnitude of targets (e.g., sample 1 has more targets than sample 2) are used to answer questions about probability (Hoemann & Ross 1982; Reyna & Brainerd in press). In risk assessment, public health officials have argued that representations of relative risk are used to answer questions about absolute risk. If the number of cases in a risk category is small (e.g., nonsmokers with lung cancer; elderly women with AIDS), increases of a few cases can produce large increases in risk relative to a baseline, even though absolute risk is low.

In each of these examples—class-inclusion reasoning, probability judgment, and risk assessment—the representations used in reasoning accurately depict the background facts: It may be true that cows outnumber horses, sample 1 has more targets than sample 2, and relative risk is large. Under certain circumstances, these representations would deliver correct responses (e.g., if sample sizes were equal in probability judgment; if the number of cases were large in risk assessment). Reasoners are more likely to be misled because representations are accurate, and because they can be used to answer a question that is similar to the one being asked. Ultimately, however, the "right way" to represent the background facts depends on the exact question.

A number of hypotheses have been advanced to explain why subjects do not answer the relevant question, including that the question is worded ambiguously (e.g., by Shipley 1979, for class inclusion, by Cohen 1981, for probability judgment, and by Henle 1962, for logical reasoning). However, when linguistic ambiguity is resolved, effects often remain: As Shipley (1979) has observed, "a majority of children persist in making errors when conventional wording is used. This suggests that there is some factor in addition to wording which interferes with correct performance in the class-inclusion task" (p. 317). The salience of inappropriate representations, too, has a small effect. For example, prominent displays increase reliance on relative magnitude in class-inclusion and probability judgment (e.g., Brainerd & Reyna 1990c; Reyna & Brainerd in press). The primary source of difficulty in these tasks, however, is in processing, which, by sending reasoning down a garden path, has repercussions for representation and retrieval.

Class-inclusion reasoning, probability judgment, risk assessment, and many other tasks, such as conditional probability, conjunction fallacy, and various deductive reasoning tasks, are subject to what has been called inclusion illusions (Reyna 1991; Reyna & Brainerd 1993, in press). Inclusion illusions occur because part-whole relationships are difficult to process, for children and for adults (Rabinowtiz et al.

1989; Spinillo & Bryant 1991; Tversky & Kahneman 1983). For any question that refers to both parts and wholes, subjects must see the same part in two ways simultaneously (as a separate part and as contributing to the whole). However, processing one level of representation tends to interfere with the other level.

For example, in class inclusion, it is difficult to view the cows as cows and as animals simultaneously. Once the cows have been mentally assigned as cows, they are no longer available to be assigned as animals, and vice versa. This creates a processing derailment: As the superordinate set mentioned in the question is considered, subsets recede, and there is no other set accessible for comparison, which produces confusion because the question mentions comparing two sets. Conversely, as processing focuses on the subset mentioned in the question, the superordinate set recedes, but this leaves the two subsets for comparison. At this point, the wording of the question does make a difference. The question seems to be about relative numerosity, and the difference in numerosity between subsets is explicit; it is represented perceptually in the stimulus display (differences between the superordinate set and the subsets are implicit). Furthermore, the wording of the question does not *contradict* assigning the subsets as the two sets to be compared. Although wording does not create the class-inclusion error, it allows it to occur.

For many problems in addition to class-inclusion, difficulties in processing part-whole relationships (i.e., nested sets) lead to an undue reliance on part-part comparisons. In logical reasoning, so-called "parentheticals" recede because part-whole comparisons are difficult to process. That is, linked elements form a salient chain, overshadowing elements outside of the chain that make up the "whole" (Falmagne 1975; Johnson-Laird 1983). For example, given premises such as *All A are B*, and *Some B are C*, reasoners mistakenly infer that *Some A are C*. However, the whole of A contains parenthetical elements that are not necessarily C. Thus, relationships between subsets of elements are used to answer deduction questions, without regard to the elements that are not linked. Similarly, in conjunction fallacy problems, elements outside the conjunction are ignored. Given a description of Linda as an outspoken person interested in social justice, subjects thought that Linda was more likely to be a feminist bank teller than to be a bank teller (Tversky & Kahneman 1983). Subjects focus on feminist bank tellers as opposed to "ordinary" bank tellers, without regard to the fact that the set of bank tellers includes those who are not feminists. Just as subjects lose track of the animals that are not cows in class inclusion and lose track of the A's that are not C's in deductive reasoning, in the conjunction fallacy, bank tellers that are not feminists are overlooked.

Subjects who, despite interference, manage to retrieve the correct reasoning principle may, nevertheless, be unable to implement the principle because of processing difficulties (Brainerd & Reyna 1990c; Reyna 1991; Reyna & Brainerd 1993). Processing can be simplified, however, by providing a notational system in which elements of parts and of wholes are distinctly represented. For example, Venn diagrams, used in syllogistic reasoning, represent subsets and more inclusive sets using a system of overlapping circles. Superordinate-set tags can be used to similar effect in class inclusion. Placing a hat on all of the animals, cows and horses, for example, reduces class-inclusion errors considerably. By providing both a retrieval cue for the correct reasoning principle (without explaining its application) and superordinate-set tags to simplify mental bookkeeping, class-inclusion errors can be virtually eliminated (Brainerd & Reyna 1990c; Reyna 1991).

Just as processing of inclusion relationships can be made simpler, so it can also be made more complex. Conditional probability judgments, for example, are more difficult than standard probability judgments (e.g., Nisbett & Ross 1980). According to Reyna and Brainerd (1993), this is because conditional probabilities involve multiple nested inclusion relationships. As we have mentioned, standard probability judgments, such as predicting which of two sample spaces is more likely to yield a target outcome, should be based on ratios (e.g., of the frequencies of target outcomes to total outcomes in each sample). However, subjects usually do not compute ratios. Instead, they estimate the relative magnitudes of targets in the two samples, a simpler and faster procedure than computing ratios (Callahan 1989; Hoemann & Ross 1982). In one recent study, for instance, only 25% of adults engaged in computation when they had the option to estimate magnitudes (Lovett & Singer 1991; see also Moore, Dixon, & Haines 1991). This procedure yields accurate solutions so long as total frequencies in each sample are equal. When samples are not equal, however, a bias to process relative magnitudes of targets should produce a pattern of errors called "denominator neglect" (Reyna & Brainerd 1993). Again, as in the other class-inclusion tasks we have considered, relative magnitude comparisons are favored because part-part relationships usurp the role of part-whole relationships.

As an illustration of denominator neglect in conditional probability judgments, Reyna and Brainerd (1993) cited systematic errors in assessing the probability of accidents on various forms of playground equipment. Certain kinds of equipment were described as more dangerous because the relative frequency of accidents on that equipment was higher. Of course, the error is failing to take the appropriate "denominators" into account, namely the frequency of use of different kinds of equipment. The probability that children were on certain equipment given that they had an accident was confused with the probability that they had an accident given that they were on certain equipment. The frequency of accidents on a certain type of equipment could be higher because children played on that equipment more often. Yet, that same equipment could be normatively less dangerous, based on the conditional probability of an accident (conditional on the probability of use).

The denominator in this example can also be called the "base rate." The failure to use base rates is a topic of longstanding interest in judgment and decision making, and it is studied in many guises in cognitive and social psychology (Kahneman et al. 1982; Nisbett & Ross 1980). According to Reyna and Brainerd (1993), neglect of base rates is a special case of denominator neglect: Because of the complexity of processing nested classes, subjects focus on numerators (e.g., joint probabilities or relative frequencies of targets, depending on the task).

Wolfe (1994) has provided specific evidence on this point. As a test of fuzzy-trace theory's predictions, Wolfe conducted three experiments concerning information seeking in solving conditional probability problems. Four types of information were provided--base rates or P(B), hit rates or P(A/B), contrapositive rates or P(A/not

B), and irrelevant information (in some conditions), and subjects were asked to estimate P(B/A). Information seeking and subjects' verbal descriptions of information (i.e., of base rates, hit rates, and contrapositives) were analyzed, and, in the last experiment, subjects were tutored about the hit rate to determine if such training would affect selection of the base rate (relative to untutored controls). The key finding in these experiments was that conversion errors—interpreting P(B/A) as P(A/B)—were associated with seeking base rates less often. Further, interventions that targeted hit rates—the locus of conversion errors—resulted in greater use of base rates.

The conditional probability of B given A, P(B/A), is equal to the joint probability, P(A&B), divided by the probability of A, P(A). The term that is confused with this, P(A/B), is equal to the same joint probability divided by the probability of B, rather than A. The latter probability, P(B), is the base rate that is ignored. Without denominators, P(B/A) is the same as P(A/B). This is the conversion error. Thus, according to fuzzy-trace theory, subjects ignore base rates because the base rate is a denominator in these problems. Overlapping inclusion relationships make processing complex (as already demonstrated in other paradigms), and subjects avoid processing complexities. Subjects avoid the nesting of probabilities within probabilities, and thus systematically mistake conversions, P (A/B), for the correct conditionals, P (B/A). As in the example about playground equipment, the two conditionals are equivalent if one erroneously ignores the denominators.

More generally, base-rate neglect as well as a host of other findings concerning relationships among probabilities indicate that subjects fail to compensate for changes in denominators (e.g., Kahneman, Slovic, & Tversky 1982). For example, Van Wallendael and Hastie (1990) showed that subjects failed to revise probability estimates for an exhaustive set of mutually exclusive hypotheses in a complementary manner. Probability estimates were consistent with the relative strength of competing hypotheses, but such estimates did not reflect changes in the overall number of hypotheses—the "denominator"—as hypotheses were eliminated. These and other results concerning relationships among probabilities can be predicted by assuming that subjects process numerators, but neglect denominators. Thus, processing complexity, having to do with inclusion relationships, provides a rationale for the regularities observed in various kinds of probability judgment.

Decision Making. Traditional theories of decision making, although they differ in many ways, share the assumption that decisions are the result of numerical processing (e.g., Lopes 1987; Payne et al. 1992). In this section, we will present data that bear on this numerical processing assumption. First, we will briefly review evidence regarding the question of whether numbers are necessary to produce decision phenomena. Then, we will present experiments that address the question of whether numbers are sufficient to produce the phenomena, and finally, whether numbers are processed at all.

Consider a choice between winning $1000 for sure versus a 50% chance of winning $2000 and a 50% chance of winning nothing. Most decision makers prefer the sure option over the gamble. This is called the certainty effect (or risk aversion). People are not uniformly risk averse, however. Preferences shift to the risky option when the outcomes involve losses. In other words, if all the outcomes in this problem were losses, most people would now choose the gamble. The shift from risk aversion to risk seeking is called a reflection or framing effect, depending on whether the same numbers are presented as losses or actual net outcomes are identical across frames, respectively (e.g., Fagley & Miller 1987).

People generally have strong preferences when they are given these kinds of choices, their preferences violate so-called normative quantitative models, and the same kinds of violations are observed when real money, and even actual life and death decisions, are involved (e.g., McNeil, Pauker, Sox, & Tversky 1982; Tversky & Kahneman 1981, 1986). According to fuzzy-trace theory, these phenomena of choice are attributable to qualitative processing (e.g., Reyna & Brainerd 1991b). Moreover, although many decisions cannot involve quantitative processing because quantitative information is unavailable, or the cost of acquiring it is prohibitive, these phenomena occur under conditions that support quantitative processing. In other words, conditions are such that memory capacity and other local information processing limitations cannot be invoked to explain qualitative thinking.

In the example and in the other problems we will discuss, there are two kinds of information: outcomes and probabilities. These outcomes and probabilities trade off so that when outcomes are higher, probabilities are lower, and vice versa. In most of the problems, competing options have equal expected value [e.g., 1.0($1000) = .5($2000) + .5 ($0)]. Although options are equal in expected value, people do not prefer the options equally; instead, as indicated, they generally have strong preferences. Therefore, preferences must be based on something other than expected value. The history of theory in decision making from the 18th century to the present can be described as, in the main, an effort to reconcile actual preferences with predictions of the expected value model.

From the beginning, the dominant theoretical strategy has been to modify the expected value model to accommodate actual behavior. For example, in order to account for preferences for sure options over gambles of equal or greater expected value, Bernoulli (1738/1954) suggested that increases in wealth produce a psychological effect that is inversely proportional to the size of one's current wealth. This implies that, as amount of money increases, the subjective value of each additional dollar decreases. Thus, a sure $1000 should be preferred over a gamble involving $2000 because larger outcomes are subject to diminishing returns.

There are a number of contemporary theories that incorporate Bernoulli's assumption about diminishing returns, including Kahneman and Tversky's (1979) prospect theory (see also Tversky & Kahneman 1992). These theories have been referred to as psychophysical because they assume that people process the numbers in decision problems much like expected value dictates, but the perception of the magnitudes of those numbers is slightly off. Other contemporary theories explain deviations from expected value by assuming that choices are a function of the variance of outcomes or of their cumulative probability (for a related discussion, see Lopes 1987).

Thus, many theorists have explained choice by proposing some modification of the expected value model. Conflicts with psychological data have been addressed by adjusting functions that transform numerical inputs or by proposing different statistics. Although newer approaches have included nonquantitative factors (such as framing choices as gains versus losses), these subjective factors are grafted onto what are, essentially, theories of numerical processing. Modern disputes center primarily on the form of the psychological equation that controls choices. That choices are a function, literally, of numerical values is generally presumed (for an alternative view, see Beach 1990).

In the first set of data we will discuss, versions of framing problems were presented in two experiments, one that we will call the Between experiment because different versions of problems were presented to different groups of subjects. The other experiment is called Within because each subject received all the versions of the problems. In the Between experiment, versions of framing problems from Tversky and Kahneman (1981) were presented (Reyna & Brainerd 1991b). Subjects received both a gain and a loss problem in counterbalanced order. Depending on their group, subjects received one of four versions of these gain and loss problems. They received the standard problems, or one of three versions in which numerical information was deleted. The numerical information was replaced by vague phrases, such as "some." So, for example, for one group, all the outcome numbers were deleted. The problem then read "If Program A is adopted, some people will be saved. If Program B is adopted, there is a 1/3 probability that some people will be saved and a 2/3 probability that nobody will be saved." For another group, probabilities were deleted, but outcome numbers were retained. Finally, for the last group, all numerical information was deleted from the standard problems. The Within experiment was like the Between experiment, except that subjects received many more problems and they received all of the versions of each problem (Reyna & Fulginiti 1992).

Results for the Between and Within experiments were highly similar. Framing differences were obtained in all conditions, namely the sure option was preferred more often in the gain frame than in the loss frame. Removing numerical information, however, did not eliminate framing effects. In fact, the largest framing effects were obtained under conditions in which all of the numerical information had been eliminated.

In the Within experiment, Reyna and Fulginiti (1992) also asked subjects to rate their confidence in their choices on a 1 to 20 scale, with 20 being the strongest. These ratings were then transformed. The ratings were given a positive sign if the sure option were chosen, and a negative sign if the gamble had been chosen. Thus, the biggest framing differential that could be observed would be 40 points, i.e., choosing the sure option with maximum confidence (+20) versus choosing the gamble with maximum confidence (–20). The ratings data indicated that subjects were highly confident about their preferences, despite the deletion of numerical information. Again, framing differences were largest when all of the numerical information had been deleted. Thus, the data from the Between and the Within experiments, both choice and ratings data, suggest that numbers are not necessary to produce framing effects.

The next logical question is whether numbers are sufficient to produce such phenomena. In the theories that have descended from the expected value model, including prospect theory, the remaining complement of the gamble—the 2/3 chance that nobody will be saved—is redundant. It literally contributes nothing to predictions about preferences. The value of a zero outcome is set at zero, and anything multiplied by zero equals zero. Therefore, for traditional theories, all of the relevant numerical information for the gamble is contained in the nonzero complement (e.g., 1/3 chance that 600 people will be saved). If only this information is processed—the zero complement is deleted—the usual framing effects should be obtained (e.g., Kahneman & Tversky 1979; Shafir, Osherson, & Smith 1989; Tversky & Kahneman 1992).

Therefore, in both a Between and Within experiment, decision makers received different partial versions of problems (Reyna & Brainerd 1991a; Fulginiti & Reyna 1993). Some subjects received only the relevant quantitative information, in order to determine whether processing this information is sufficient to produce framing effects. Subjects also received problems in which this "relevant" numerical information was omitted, but the zero complement was retained. For these zero-complement problems, the gamble simply stated, for example, that there is a 2/3 chance that nobody will be saved. Although the zero complement is superfluous quantitatively, according to fuzzy-trace theory, it contains the pivotal qualitative contrast that is responsible for framing effects (e.g., that nobody might be saved, as opposed to some saved for sure; see the earlier discussion on framing).

The results were quite clear. In the Between experiment, when subjects were given the relevant numbers, which ought to be sufficient to produce framing effects, there were no framing effects. However, when subjects were presented with the zero-complements—the part of the gamble that did not contain precise numerical information—they showed large framing effects. Similar results were obtained in the Within experiment. Multiple decisions were presented in the Within experiment, some involving money others involving lives, and all showed this pattern. Again, the rating data were consistent with the choice data. Subjects's confidence ratings were close to indifference when the relevant numbers were presented, but they had strong preferences, consistent with framing, when that information was omitted.

The logic of these manipulations, presenting parts of the gambles, is to induce subjects to selectively process the information that is presented. The idea is to focus processing on certain problem components to determine if processing those components is sufficient to produce framing. As we have seen, processing just the relevant numbers does not seem to be sufficient to produce framing. On the other hand, processing the zero complement produces large framing effects, as predicted by fuzzy-trace theory. One might argue, however, that problems with missing parts are ambiguous. Subjects might "fill in" partial problems, rather than responding just to the presented information. There are two sets of evidence that speak to this ambiguity explanation.

First, in the Within experiment just discussed, partial problems were randomly interspersed among many complete versions of the same problems. Thus, as subjects proceeded through the experiment, uncertainty about missing elements was reduced. In order to test the hypothesis that selective processing effects were due to ambiguity, choices for the same partial problems at various stages in the experiment were compared. Reyna and Fulginiti (1992) analyzed the proportion of subjects choosing the sure option across six blocks of trials. (As a baseline for comparison, it should be noted that preferences for the complete gain and loss problems were stable across all six blocks.) For partial problems with the relevant numerical information, the standard framing pattern was actually reversed in the early blocks. However, after multiple exposures to complete problems—which specified the missing zero complement—the choices crossed over. The signed confidence data showed the same crossover. Relevant numbers not only failed to produce the usual framing effects in the early blocks, significant differences in the wrong direction were found.

In contrast, for partial problems with just the zero complement, choices did not change across trials. That is, knowledge about omitted numerical information failed to affect choices. Paradoxically, partial problems in which all the relevant quantities were already specified were affected by exposure to supposedly irrelevant information, but zero-complement problems that lacked precise information were not affected by exposure to supposedly relevant numerical information.

The other evidence against an ambiguity explanation for results with partial problems is straightforward. Another Between experiment was conducted with subjects who received multiple decision problems (Reyna 1992b). Half of the subjects received zero complement problems only, and the other half received problems with just the relevant numerical information. For both groups, however, gambles were disambiguated. For each problem, the missing information was stated above the options. For example, in the relevant-numbers group, for a problem pitting a sure win of $1000 against a 50% chance to win $2000, a preliminary sentence stated that "For Option B, if you do not win the $2000, you neither win nor lose anything." One problem was presented per page with the disambiguating information prominently displayed at the top, attention to "every word" was stressed in both oral and written instructions, and subjects indicated awareness of the disambiguating information during debriefing.

In effect, then, both groups of subjects received identical information—only the focus of processing was shifted. Although there was no ambiguity about information that was omitted in the body of the problem (because that information was stated explicitly), selectively focusing on parts of the gamble yielded the same pattern discussed earlier: no relationship between processing relevant numbers and framing, but large framing effects when subjects focused on the zero complements. Again, rated confidence in choices was high when subjects processed the zero complements, but low when the relevant numbers were highlighted.

In toto, the evidence suggests that processing relevant numbers is not sufficient to produce framing. This leads us to our last question, Are numbers processed at all? Do subjects fail to process relevant numbers because of memorial or computational deficits? Fulginiti and Reyna (1993) presented subjects with three kinds of problems to determine whether they were sensitive to numerical information. These were standard problems, converging problems, and diverging problems. Expected values were not equal for converging and diverging problems. For converging problems, the alternative that was not favored based on framing was made slightly superior in expected value. So, for example, subjects had to choose between saving 199 people for sure versus a 1/3 chance of saving 603 people (and a 2/3 chance of saving nobody). Thus, the gamble was favored (had a slight quantitative advantage) in the gain frame and the sure thing was favored in the loss frame. Diverging problems were exactly the opposite. The sure option was favored in the gain frame, and the gamble was favored in the loss frame. Options in standard problems had equal expected value. Subjects received all three types of problems.

Subjects were indeed sensitive to quantitative differences. Framing effects were diminished when they conflicted with quantitative differences, as in converging problems. And, framing effects were augmented by compatible quantitative differences. Signed confidence ratings showed the same patterns. Interestingly, the same subjects who showed framing effects when expected values were equal, responded to small inequalities in expected values. Thus, it appears that subjects show framing effects *in spite of* their ability to process expected values.

In summary, these experiments on decision making suggest that numbers are not necessary to produce well known decision phenomena, nor are they sufficient to produce those phenomena. Numbers were not necessary because eliminating them failed to eliminate standard effects. Selectively processing supposedly crucial numerical information, however, did eliminate these effects. Numbers were not sufficient because partial problems with only relevant numerical information failed to produce standard framing effects, whereas partial problems with zero complements produced large effects. As long as partial information remained the focus of processing, the same effects were obtained, even when subjects were informed about missing information (indicating that inferences about missing information failed to account for these effects). Finally, although preferences were not attributable to numerical processing, subjects responded to subtle differences in expected values, indicating that they are capable of processing that numerical information. All of this implies that decision phenomena are the result of qualitative reasoning, as opposed to the psychophysics of numbers or their statistical properties, in spite of the fact that decision makers encode and can process quantitative information correctly.

Cognitive Competence. The finding that decision makers can process quantitative information correctly speaks to their competence, but how should we evaluate the quality of their thinking overall? In Piagetian theory, the major source of errors in reasoning is lack of logical competence. Children fail transitivity, probability, class-inclusion, and other tasks because they fail to understand the logic of relationships implicit in problem facts. Information-processing theories explain reasoning errors in terms of local breakdowns in memory (e.g., effects of memory "load"). Fuzzy-trace theory shifts the explanatory focus away from logical deficits or memory for information to the ability to recognize and represent the global patterns embedded in that information. The theory departs sharply from traditional approaches in assigning a central role to gist in advanced reasoning. In this view, a "fuzzy-processing preference" represents a system-wide adaptation to the limits of information processing, a means of avoiding systematic errors caused by poor verbatim memory (Reyna & Brainerd 1992).

In fuzzy-trace theory, therefore, thinking is seen as fundamentally intuitive. This contrasts with the Piagetian conception of thinking as the rigid application of logical rules to premise-like inputs. It also contrasts with the mind-as-computer metaphor of information-processing, which emphasizes elaborate knowledge representations and identifies precision and quantification with accuracy. Not only does the conception of thinking differ in fuzzy-trace theory, so do ideas about development. The fuzzy-processing preference is ascribed to mature reasoners, and represents a flexible and adaptive approach to reasoning that, overall, has the effect of reducing errors. This is the opposite of the traditional view that development progresses away from intuition towards greater logic or computation.

Errors do occur, however, and there are two contrasting developmental trends in errors. With respect to the classic cognitive developmental paradigms, errors decline with age. Computational skills clearly improve during childhood and, for some tasks, must be invoked (Moore, Dixon, & Haines 1991). On the other hand, certain judgmental and decision-making biases increase with age. For example, Jacobs and Potenza (1991) found that young children's probability judgments were not biased by the "representativeness heuristic," but use of the heuristic increased with age, and adults' judgments were inferior to those of children. Similarly, older children are more likely to use noncompensatory reasoning strategies in decision making under certainty (i.e., that did not involve risk), compared to younger children (see Klayman 1985). Reyna and Ellis (in press) found that the framing bias, too, increased with age, and like probability judgment and decision making under certainty, risky decision making departed increasingly from the ideal quantitative model (see also Reyna & Brainerd 1993).

Moreover, these contradictions are all the more puzzling because some of the same concepts are involved in the cognitive development and in the judgment and decision making literatures. For example, based on research in cognitive development, one would conclude that probability judgment improves as children get older (Reyna & Brainerd in press). The implication would be that once those same children became adults, they would perform at least as well, if not better than, young children. However, if we define "better" as processing ratios of frequencies, that implication is false. Adults are likely to use relative magnitude of targets rather than ratios to judge probability (although they are capable of computing the correct ratios). Adults are also more likely to reject appropriate quantitative information in a biasing context than younger children (Jacobs & Potenza 1991); and, they are more likely to make intuitive rather than quantitative judgments when the latter are optional (Lovett & Singer 1991). Finally, adults' errors in probability judgment have been demonstrated in numerous studies on "heuristics and biases" (Kahneman, Slovic, & Tversky 1982). It is not just that adult cognition is not quite as advanced as it should be. The biases research and the cognitive developmental research imply opposite views of adult competence in judgment and decision making.

One might argue, that children and adults select from a menu of rules (Siegler 1988) depending on the task, and that the variability across tasks cannot be spanned by domain-general assumptions. Alternatively, one might argue, with Piaget, that variability proves that formal competence has not been acquired because such competence, by definition, spans tasks that differ superficially (e.g., Chapman & Lindenberger 1992). Of course, by this standard, one would be forced to conclude that adults lack the probability concept, as well as many other concepts (Nisbett & Ross 1980). Indeed, experts have been shown to exhibit the kinds of biases in probability judgment and in decision making that we have discussed (Kahneman et al. 1982; Reyna & Brainerd 1993). By varying the task slightly, adults can even be made to fail tests that tap seemingly elementary concepts such as conservation (e.g., Winer & McGlone 1993).

Fuzzy-trace theory would account for this variability in performance for children and adults by distinguishing between competence—knowledge of a concept and how to apply it (albeit under limited circumstances)—and various other potential sources of errors. We would argue that adults generally have quantitative competence in these tasks. They are able to process information in a quantitative fashion, but they tend not to engage in such processing unless the task requires it. Young children's computational skills are more limited than those of adults. However, for many tasks, young children know the relevant concept, and can sometimes demonstrate that competence.

For example, in class-inclusion, children answer questions such as "Are there more people in Tucson or in Arizona?" correctly, so long as displays of numerical information are not presented. In the standard class-inclusion task, asking children to wait 15 s before responding improves performance (Winer 1980). In transitive inference, within a single session, performance improves with mere practice. In probability judgment, sensitive techniques, such as functional measurement, reveal advanced competence among first graders (Acredolo, O'Connor, Banks, & Horobin 1989). Retrieval prompts, such as reminding children of problem facts or cuing relevant reasoning principles, are sometimes sufficient to eliminate errors (Brainerd 1981; Brainerd & Reyna 1990c).

Correct reasoning, however, involves more than competence. It also involves encoding the appropriate gist in problem information, inhibiting interference from irrelevant verbatim details, editing out irrelevant gists, retrieving the appropriate principle in context, and correctly implementing that principle (i.e., correctly applying the principle to the gist representation of the problem). Each of these components has been shown to make independent contributions to successful reasoning (e.g., Brainerd 1983; Brainerd & Reyna 1990b, 1993; Reyna 1991; Wilkinson 1982).

With development, children become less subject to interference from both verbatim details and from competing gists, and they also retrieve and process relevant principles more reliably (Brainerd & Reyna 1993; Reyna 1991; Reyna & Brainerd 1991a). Encoding the appropriate gist does not seem to be the source of much developmental variation, especially after first grade. Commonly, an array of gists are activated in a specific problem-solving context, despite the fact that most are not used. However, children may focus on salient relationships that are inappropriate to the question at hand.

Adults, too, fall victim to errors of representation, retrieval, and processing. Although such errors are less likely in adults than in children, they are not fundamentally different (Reyna & Brainerd 1993). Regarding representations, adults are sometimes misled by salient patterns into responding to the wrong gist (Tversky & Kahneman 1981, 1983). The same psychological processes that make representation and retrieval more reliable in adults, make biases more reliable. With experience, adults come to recognize familiar patterns quickly, and may leap to retrieval of well-known, but inappropriate, principles. Therefore, manipulating the cues in a problem produces systematic errors in reasoning (e.g., Kahneman et al. 1992; Reyna & Brainerd 1991a). As inclusion illusions (e.g., denominator neglect), false recognition errors, and a host of other effects demonstrate, adults are less subject to interference, but they are not immune. Therefore, although development involves increasing resistance to interference, and greater reliance on the gist, mature processing has predictable pitfalls.

Rather than characterizing reasoning as either correct or incorrect, or understanding as either present or absent, fuzzy-trace theory identifies different levels of understanding. In order of sophistication, these range from failure to know a reasoning principle, to failure to appreciate the relevance of that principle in varied contexts, to failure to recognize and correct misdirections in reasoning that are due to interference (Reyna 1995; Reyna & Brainerd 1990, 1991a). Does the fact that these failures occur in mature reasoners, that some biases actually increase with age, and that reasoning is mostly intuitive (rather than logical or computational), indicate that human reasoning is irrational?

Rationality. To say that adults possess underlying competence, and have had it from early childhood, is not necessarily to claim that their reasoning is rational. There are several distinct definitions of rationality that have been proposed, and that bear on the issue of whether intuitive reasoning can be classified as rational (Reyna & Brainerd in press).

The question of rationality is central to developmental research primarily because theorists have assumed that cognitive development is progress toward rationality (e.g., Piaget 1949; Werner 1948). As we have seen, however, this places us on the horns of a dilemma: We must accept experimental data showing that adult reasoning is flawed, but we must also somehow explain the course of development. As it turns out, by carefully exploring definitions of rationality, this conflict can be reconciled.

Tversky and Kahneman (e.g., 1983) have used the term "intuitive" to refer to the kind of natural processing that leads to biases and errors in probability judgment (e.g., the conjunction fallacy). In this usage, they can be grouped roughly with Piaget for whom intuition was, by its nature, the absence of rationality (i.e., the absence of logic). However, Tversky and Kahneman (and others) have discussed at least three kinds of rationality, including correspondence to extensional reality, correspondence to formal rules (e.g., principles of inferential statistics), and various forms of internal coherence among judgments (e.g., the invariance axiom).

Human behavior, including judgments and choices, are considered rational if they are consistent with reality, for example, whether preferences between alternatives map onto *actual* differences in outcomes (Tversky & Kahneman 1981, 1986). To oversimplify somewhat, assuming that one would prefer to be richer, it is irrational to pick an option that makes one poorer (compared to its alternative). Although gist does not represent literal reality, there is nothing intrinsic to gist-based representations that would entail any contradiction with reality. Moreover, the fuzzy-processing preference is generally constrained by the level of precision demanded in a task. For example, when alternatives begin to diverge in expected value, subjects do note the nonequivalence of options and shift responses. Further, processing fuzzy, or gist-level, representations tends to minimize reasoning errors (Estes 1980; Reyna & Brainerd 1991a). Taken together, these considerations suggest that intuitive processing allows the individual to maximize overall performance by trading off precision and simplicity.

For many judgments, it is difficult to determine what reality is. In these cases, the second criterion of rationality—correspondence to formal rules—is used. In Tversky and Kahneman's (1983) well-known conjunction fallacy problem, the Linda problem, there is no way to know the probability that Linda (or someone fitting her description) is a feminist. Logic, mathematics, and other formal systems, however, allow us to place constraints on relations among judgments without knowing the exact state of reality. Biases and errors, then, are judged with respect to specific canons of logic, Bayesian probability theory, and other formal systems. If the relevance of those rules is not granted, then subjects' responses are not necessarily errors. Although reason has historically been defined with reference to logic and mathematics, it is clear that human reasoning often does not adhere to these formal systems. However, intuitive reasoning is not necessarily prone to error, quite the contrary in practice. Therefore, failure to be logical or quantitative does not seem to be sufficient to characterize reasoning as irrational.

The most basic criterion of rationality does not require any particular definition of reason. Nor does it require exact knowledge of reality. Whatever the basis for subjects' responses, the minimal condition that those responses must satisfy is invariance (Tversky & Kahneman 1986). Judgments of the same information by the same individual should at least (barring statistical error, the passage of time, or changes in circumstances) be consistent with one another. If judgments are inconsistent, it is difficult to claim that there is some rational basis for those judgments.

Framing effects, for example, violate invariance; most people prefer a sure gain when it is pitted against a gamble of equal expected value, but if the choices involve losses, most people prefer the gamble. So, superficial changes in wording—even when the same outcomes will be experienced—cause subjects to shift from risk aversion to risk seeking. This shift is taken as evidence of irrationality (e.g., Tversky & Kahneman 1986).

By this most fundamental criterion, gist is key to rationality. By operating on the underlying gist of information, rather than on verbatim details, reasoning can be invariant across superficially different problems. But, is it? Developmentally, according to fuzzy-trace theory, reasoning becomes increasingly invariant. That is, correct reasoning is applied to a wider range of tasks that abstractly instantiate the same concepts (Brainerd & Reyna 1990b). Age-related improvements in resistance to interference, and in retrieval and processing foster the development of invariance—the ability to consistently apply correct principles to appropriate tasks. However, as we have seen, reasoning never becomes entirely invariant. There seem to be some predictable pitfalls in reasoning that are created by the need to trade off two competing aims, the need to use the same information to answer different questions (cognitive flexibility) and the need to respond to superficially different situations similarly (cognitive consistency).

DEVELOPMENT OF FORGETTING

Developmental research on forgetting from long-term memory has a rather confusing history. We say this because there has been a deep disparity between what common-sense expects and what laboratory studies have usually found. Naturally, common-sense expects that forgetting will be profound early in life (e.g., the idea of infantile amnesia), will then decline between early childhood and young adulthood, and will eventually pick up again late in life. Binet's (1900) book, *La suggestibilite*, is one of the earliest scientific works in which the forgetfulness of young children figured prominently. Binet claimed that young children's memories were so labile that leading questions and other forms of suggestion would produce "recollections" of events that had never been experienced. Outside the realm of science, much the same notion is enshrined in legal practices governing the admissability of children's testimony. Until very recently, the courts regarded children's memories as too ephemeral to allow for accurate testimony. This ruled out the prosecution of crimes, especially sexual abuse, in which children were the only witnesses and there was no corroborative physical evidence.

Laboratory studies of the development of forgetting provided a completely different picture. Between the early 1970s and the mid-1980s, a series of studies appeared that all showed the same thing: no age variability in forgetting rates (for a review, see Brainerd, Kingma, & Howe 1985). These studies made use of the standard procedure for studying normal (i.e., noninduced) forgetting—the Ebbinghaus retention paradigm. In this procedure, subjects are first exposed to some target information (words, pictures, stories, films), following which memory for some or all of it is tested. A forgetting interval—typically, a few hours or days—then ensues, at the end of which more memory tests are administered. The decline in performance between the immediate and delayed tests is the measure of forgetting. Lehman, Mikesell, and Doherty summarized what such studies had shown up to the mid-1980s: "information is not lost more rapidly by children than by adults. Forgetting rates ... were invariant from childhood to young adulthood" (1985, p. 27).

This conflict between common-sense and what had been found in the laboratory was the original impetus for our own studies of the development of forgetting. To date, those studies have focused on three main questions: Can age differences in forgetting rates be demonstrated under appropriate conditions? Are age differences in forgetting rates epiphenomena of correlated age differences in learning rates? What theoretical mechanisms are responsible for the development of forgetting? The present section is organized around these questions.

Does Forgetting Develop? As of the mid-1980s, then, there seemed to be compelling evidence that forgetting rates do not vary between the preschool years and the early 20s. At about this time, however, it was suggested that the studies that had produced this outcome incorporated design variables that, with the wisdom of hindsight, could have masked age variability (Brainerd et al. 1985). Eventually, three candidate variables were identified (for a review, see Brainerd et al. 1990): Recognition insensitivity, floor effects, and levels-of-learning confounds.

Recognition insensitivity refers to a difference in the power of recognition and recall tests to detect age variability. If the objective is to detect age variability in some memory process, it goes without saying that tests that are maximally sensitive to developmental change should be used. In most of the studies that had produced the no-development finding (e.g., Fajnsztejn-Pollack 1973), both the immediate and delayed tests involved some type of recognition. As is well known, however, performance on recognition tests displays limited age variability (e.g., Ornstein & Corsale 1979). The use of such tests in retention designs therefore stacks the deck against observing developmental trends in forgetting rates.

The floor effect problem is even simpler. An obvious precondition for observing age variability in forgetting is that there should be a significant decline in performance between the immediate and delayed tests. This, after all, is the operational definition of forgetting. So, there must be statistically significant forgetting before there can be statistically significant age changes in forgetting. For a confluence of reasons, such as the use of short forgetting intervals (e.g., 1 day) and very memorable items (e.g., pictures), amounts of forgetting were very small in most studies that produced the no-development finding. For instance, Merriman, Azmita, and Perlmutter (1988) reported a study of declines in picture recognition across a 1 day interval for 3-, 4-, and 6-year-olds. Since the average decline was only 5%, there was little room for age variability. Likewise, Morrison, Haith, and Kagan (1980) reported two developmental studies of forgetting in which most conditions failed to produce statistically significant forgetting, regardless of age level.

The third problem, stages-of-learning confounds, is the most complicated one. In the adult memory literature, a good deal of attention has been devoted to the question of how to control for the fact that between-conditions differences in forgetting rates may be due to correlated differences in initial level of learning (e.g., Underwood 1964). This question is of concern in developmental studies of forgetting because, naturally, there will be developmental differences in learning at the end of an acquisition session. Other things being equal, older children (and adults) will learn more about target material than younger children will, and this will be reflected in performance differences on the immediate memory tests. Moreover, because learning curves are negatively accelerated, this built-in age difference in levels of learning is necessarily maximal when subjects receive only a single exposure to the material (Brainerd et al. 1985).

Built-in age differences in initial levels of learning can mask age differences in the subsequent amounts of forgetting for simple statistical reasons (cf. Brainerd, in press a). Suppose that younger children and older children study a set of 100 memory targets and then receive a recall test. Suppose that the younger children recall 25 targets and the older children recall 50 targets. Two weeks later, all the children receive a second recall test. Suppose that during the two-week interval, the forgetting rate for younger children is .8 (80% of the targets that were recalled on the immediate test will not be recalled on the delayed test), and the forgetting rate for older children is .4 (40% of the targets that were recalled on the immediate test will not be recalled on the delayed test). Because of the difference in initial learning level, however, the number of forgetten items will be the same at both age levels: $25 \times .8 = 20$ forgotten items for the younger children and $50 \times .4 = 20$ forgotten items for the older children.

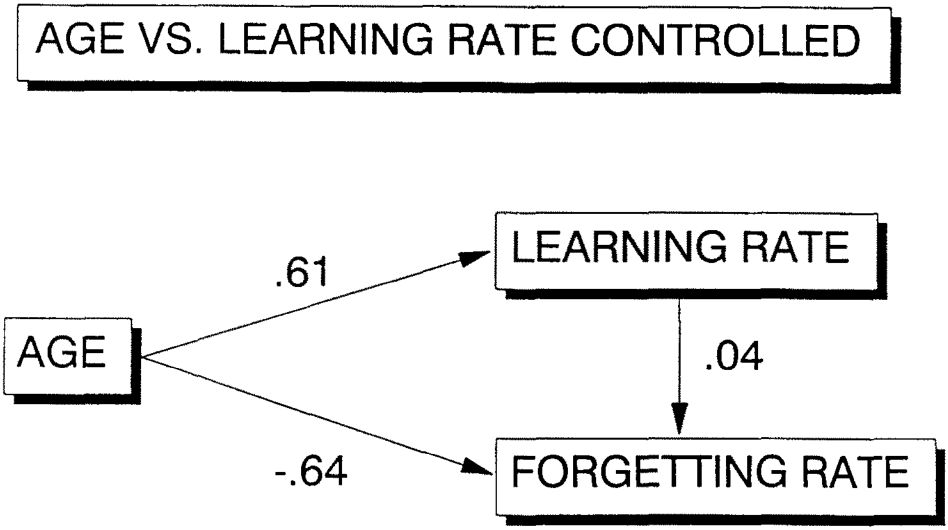
In light of such problems, we conducted a series of new studies using modified designs (e.g., Brainerd et al. 1990; Brainerd & Reyna 1991). The recognition insensitivity problem was dealt with by switching to some form of recall, the floor effects problem was dealt with by using longer forgetting intervals (e.g, 2 weeks rather than 1 day), and the levels-of-learning problem was dealt with by requiring subjects of all age levels to learn the material to a perfect-recall criterion. Our studies have all shown the same thing: common-sense beliefs are confirmed. That is, forgetting rates are highest during early childhood, decline steadily between childhood and adolescence (especially between early and late childhood), and eventually increase again late in life (e.g., Howe & Hunter 1986).

Are Differences in Forgetting Just Age Differences in Learning? It might be thought that the question of whether forgetting rates vary with age has now been answered. Not so, unfortunately. It turns out that one of the three modifications, the one that dealt with levels-of-learning confounds, introduces a potential confound of its own: age-learning rate correlations.

In classical theories of forgetting (e.g., McGeogh 1942; Underwood 1954), it is assumed that individual differences in learning rates determine individual differences in forgetting rates—specifically, that subjects who learn material faster forget it more slowly. This leads to an obvious worry about criterion learning. Although this technique eliminates age differences in level of initial learning, by requiring subjects of all age levels to meet the same criterion, it does so only at the expense of introducing age differences in the rate at which the criterion is reached. If, for instance, younger and older children first learn a list of concrete nouns to a perfect-recall criterion, it might take an average of ten trials for younger children but only five trials for older children. This could account for subsequent age differences in the rate at which the nouns are forgotten.

We tested this hypothesis in two experiments and found no support for it (Brainerd & Reyna 1990a). In both experiments, 7- and 11-year-olds first learned noun lists to criterion under standard free-recall conditions. After a two-week

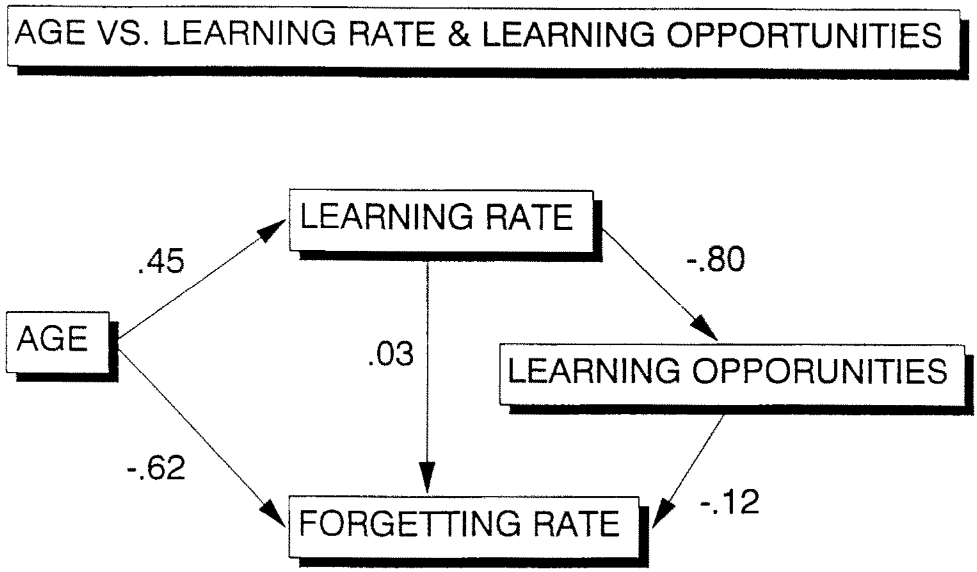
FIGURE 3
Combined results for path analyses reported by Brainerd and Reyna (1990a).



forgetting interval, they returned for a series of retention tests, which consisted of four more cycles of free recall without further opportunities to study the words. Since performance had been constrained to be perfect two weeks earlier, the error rate during the retention session was the measure of forgetting. As in prior experiments using this design (e.g., Brainerd et al. 1990), the younger children forgot more than the older children. We then used path analysis to determine whether this outcome was due to correlated age differences in learning rate. Three variables were entered in the analysis: age (in months), forgetting rate (error rate at retention), and learning rate (trial of last error, total errors to criterion at acquisition). The pooled results for the two experiments are displayed in Figure 3. As can be seen, there was an age - - -> forgetting causal path (forgetting rate decreased with age), an age - - -> learning causal path (learning rate increased with age), but no learning - - -> forgetting causal path. Thus, developmental trends in forgetting could not be explained as by-products of correlated age differences in learning rates.

Although these results were encouraging, there was another possible confound that might explain why there is no learning rate-forgetting relationship in Figure 3: opportunities to learn. Measures of learning rate will automatically correlate with opportunities to learn (e.g., number of learning trials) in any criterion design; the faster the learning, the fewer exposures to the target material before reaching criterion. This would not necessarily be problematical were it not for the fact there is also an potential confounding relationship between forgetting and opportunities to learn: The tendency to forget material may decrease as the number of exposures to it increases. If so, it will be hard to detect a learning rate-forgetting relationship

FIGURE 4
Combined results for path analyses reported by Brainerd and Reyna (in press b).



because the learning rate-learning opportunities and the learning opportunities-forgetting relationships will mask it. Specifically, even if slower learning means faster forgetting, the additional learning opportunities that must be given to slow learners may compensate for this difference. What is needed, obviously, are some further path analyses in which the learning rate variable is added to the three variables in Figure 3.

We have conducted analyses of this sort (Brainerd & Reyna in press b), and they failed to show that Brainerd and Reyna's (1990a) results were artifacts of failures to control learning opportunities. Again, there were two studies in which 7- and 11-year-olds first learned lists of words to criterion under free-recall conditions and then returned two weeks later for a series of retention tests. The path analyses were like those in Figure 3, except that a measure of learning opportunities (trials to criterion) was added to the original three variables. The pooled results for the two studies appear in Figure 4. As can be seen, although there was a strong learning rate-learning opportunities relationship, there was neither a learning rate-forgetting relationship nor a learning opportunities-forgetting relationship. The important conclusion, once again, is that there is true age variability in forgetting rates that is not contaminated by correlated age variability in learning rates.

Why Does Forgetting Develop? Since epiphenomenal explanations do not work, the question becomes, What does cause age variability in forgetting rates? Our own attempts to answer this question have focused on two theoretical distinctions. The first, which has long been a core idea in memory research, is the distinction between storage-based and retrieval-based forgetting. The other, which is central to fuzzy-trace theory, is the distinction between verbatim-based and gist-based storage failures.

1. Storage failure versus retrieval failure. Perhaps the most fundamental distinction in mainstream theories of forgetting is between forgetting in the sense of progressive loss of memory traces (storage failure) and forgetting in the sense of progressive loss of the ability to access otherwise intact traces (retrieval failure). Historically, opinions have varied as to the relative contributions of these two processes (Loftus & Loftus 1980). At the beginning of the century, the predominant view was that forgetting was due to trace decay, trace reorganization, and other forms of storage failure. The opposite view, that traces remain intact following storage and that forgetting is synonymous with retrieval failure, became prevalent in the 1930s with the advent of interference theory (e.g., McGeoch 1932).

This retrieval-failure hypothesis has remained the predominant interpretation (Brainerd et al. 1990). Loftus and Loftus (1980) noted that, nowadays, two types of experiments are most often cited in connection with this hypothesis. The first consists of studies (e.g., Penfield & Roberts 1959) that appeared to show that weak electrical stimulation of the cortex can produce retrieval of vivid recollections of events that (putatively) were experienced many years before. The second consists of studies (e.g., Ballard 1913) of reminiscence, the familiar phenomenon of being able to recollect something at Time 2 that could not be recollected at Time 1. With both types of experiments, the favored interpretation is that traces have remained intact since they were stored, but it has become difficult to retrieve them.

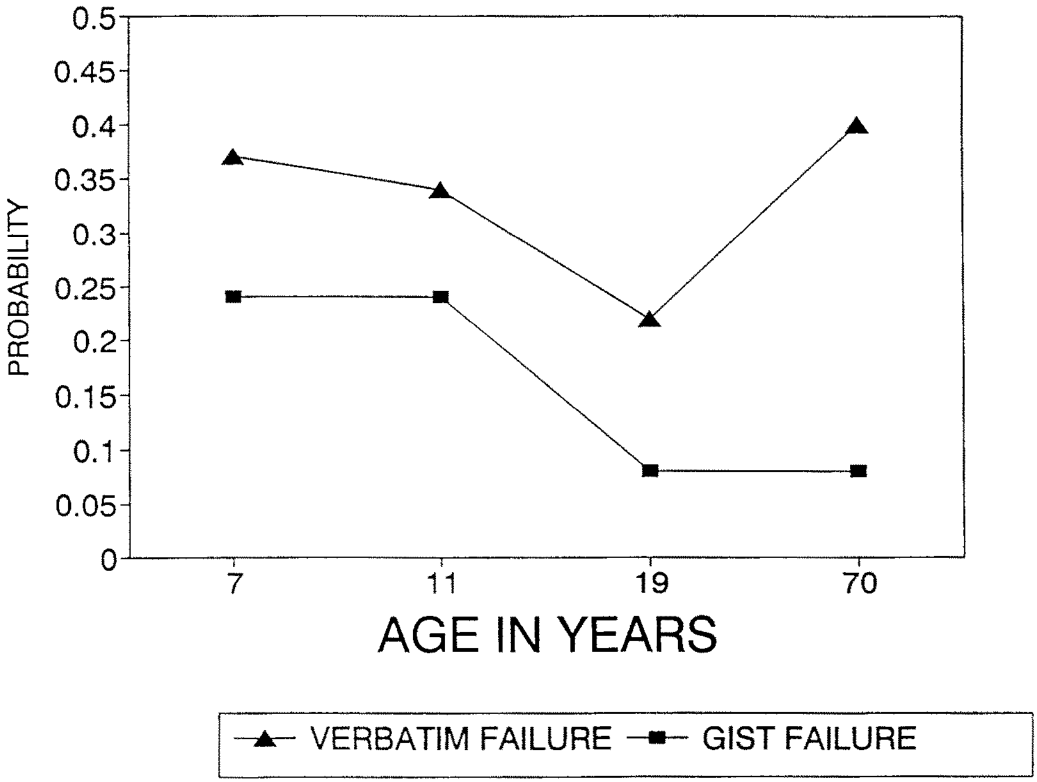
We have studied the relative contributions of storage and retrieval failures to forgetting within and between age levels using mathematical models whose parameters provide ratio-scale measures of these processes. These models evolved from some earlier ones that were developed to measure the rates at which subjects initially store information and learn how to retrieve it (Brainerd 1985a, 1985b). Without going into the mathematical details, the forgetting models contain some parameters that measure the rate at which traces disintegrate over forgetting intervals and some other parameters that measure the rate at which the ability to access still-intact traces is lost (Brainerd 1986; Howe & Brainerd 1989). These models have now been applied in many forgetting experiments with preschoolers, elementary schoolers, young adults, and the aged (Brainerd & Reyna 1991; Brainerd et al. 1990; Howe 1991; Howe, Courage, & Bryant-Brown 1993; Howe, Kelland, Bryant-Brown, & Clark 1992; Marche 1993).

From the standpoint of the retrieval-failure hypothesis, the results have been quite surprising. Although this hypothesis assumes that traces remain intact, substantial amounts of storage failure have been measured in all of these experiments. Even more surprisingly, throughout the life span in nearly all experimental conditions, the storage-failure rate has far exceeded the retrieval-failure rate. In fact, for certain types of designs, the retrieval-failure rate has approached zero (cf. Brainerd et al. 1990). A final important finding has been that storage failure often contributes more to age variability in forgetting rates than retrieval failure does. For example, Brainerd et al. (1990) summarized the developmental pattern that emerged from their series of experiments as follows: "the development of forgetting does not appear to be dominated by retrieval failure. There were no age comparisons where this was true" (p. 77).

2. Verbatim failure versus gist failure. If storage failure is the principal source of forgetting, it is important to unpack this concept and investigate failure rates for different types of memories. In fuzzy-trace theory, of course, interest attaches to the rates of both verbatim and gist failure. We have investigated this question with the aid of some further mathematical models of forgetting (Brainerd et al. 1994).

The designs of most of our developmental studies of forgetting, as well as the nature of the memory targets that have been administered, tend to ensure that forgetting will be more a matter of verbatim failure than of gist failure. As is the case in most child memory research, the memory targets always consisted of familiar materials (concrete nouns, pictures, sentences describing everyday events) that instantiated concepts that even the youngest children could presumably access

FIGURE 5
Life-span patterns of age variability in verbatim and gist forgetting rates.



in their long-term memories. Given the targets BLUE, COW, and CHAIR, for instance, even preschoolers will identify BLUE as a color, COW as an animal, and CHAIR as a type of furniture (Bjorklund 1989). Further, since initial learning was to a perfect-recall criterion in our designs, younger children were given many opportunities to access such concepts during the course of learning. Thus, the designs and the memory targets tended to eliminate the possibility of failures to store gist during learning. If gist traces are easier to retain than verbatim traces, this means that subsequent forgetting should be more a matter of verbatim failure than of gist failure at all age levels.

That is what we have found in a series of model-based decompositions of the age trends in our forgetting experiments. The decompositions rely on some mathematical models that measure the probability that a recall error at the start of a delayed test is a verbatim-memory failure versus a gist-memory failure (Brainerd et al. 1994). They are elementary Markov processes that consist of a state V (verbatim-memory failure), a state G (gist-memory failure), and a state R (no memory failure). These models allow us to assess both the rate of verbatim forgetting and the rate of gist forgetting by calculating how many targets occupy state V and how many occupy state G at the start of a delayed test.

Measurement of the respective rates of verbatim and gist forgetting for our data sets has usually produced two results. First, at all age levels, verbatim-memory failures tend to contribute more to recall errors on delayed tests than gist-memory failures; most forgotten targets occupy state V rather than state G. Second, both the verbatim-failure rate and the gist-failure rate tend to vary with age. The exact pattern of variability depends on the age range that is being studied and on various design considerations (Brainerd et al. 1994). However, an interesting illustrative pattern is shown in Figure 5, where verbatim and gist failure rates are plotted by age for some life-span data sets in Brainerd et al. (1990, Experiments 3 and 4). In these experiments, 7-, 11-, 19-, and 70-year-olds received a delayed recall test either 1 or 2 weeks after learning sets of pictures or concrete nouns to a perfect-recall criterion. As can be seen, the life-span trend for the verbatim forgetting rate is that it declined between age 7 and age 19, and then it increased between age 19 and 70. The life-span trend for the gist-forgetting rate is that it also declined between age 7 and age 19, but it did not vary between age 19 and age 70.

THE DEVELOPMENT OF RETRIEVAL: COGNITIVE TRIAGE

An assumption that fuzzy-trace theory shares with some other recent theories (e.g., Dempster 1992; various contributors to Dempster & Brainerd 1995) is that being able to resist interference is an important factor in memory and reasoning. We now take up a counterintuitive phenomenon that illustrates this contention, cognitive triage, in that it is concerned with the role of a specific type of proactive interference (output interference) in retrieval and in its development. We call it cognitive triage because, as will be seen, it is reminiscent of the battlefield surgical technique of dealing first with the most acute injuries. Cognitive triage is also another instance of the fertility of the verbatim-gist distinction. We consider theoretical ideas about recall first and then summarize major findings from cognitive triage experiments.

Recall Order: Theory. Free-recall tasks are laboratory analogues of those everyday situations in which we retrieve sequences of episodically-related events from our long-term memories. (What did you do at school today? Who came to your birthday party?) In the standard task, subjects study a series of memory targets (words, pictures, etc.) and then recall as many of them as possible in whatever order they come to mind. Some targets will be harder to retrieve than others, and, across a series of study-recall cycles, error rates will be higher for those items. The question that will concern us is, What is the relationship between the *retrieval difficulty* of individual targets (as measured by error rates) and the *order* in which they are read out during free recall?

The answer given by most memory theories is that the order will be easier - - -> harder. The basis for this prediction varies. One of the earliest is Marbe's Law (1901), which posited that the activation speed of a memory trace decreases logarithmically with the strength of that trace. The easier - - -> harder ordering also falls out of several modern theories, including associative network theories (e.g., Anderson & Bower 1973), the resources hypothesis (e.g., Bjorklund & Muir 1988), and the search-of-associative-memory (SAM) model (e.g., Shiffrin, Ratcliff, & Clark 1990). Associative network theories deliver this prediction because they assume that the memory nodes where targets are stored are connected via associative bonds, that retrieval difficulty depends on the sum of activation at a target's node, and that activation spreads through these networks in an easier - - -> harder direction. The resources hypothesis, on the other hand, postulates a limited-capacity pool of mental energy (the processing resource) that must be consumed to support recall. Since easier targets tax this resource less than harder ones, the easier - - -> harder ordering maximizes recall because it has the slowest possible rate of resource consumption per item recalled. Last, SAM treats recall as a process whereby subjects sample memory images and read out the information stored in them. Images are sampled in order of their strength (i.e., stronger - - -> weaker), and strength also determines overall error rate (i.e., stronger = easier).

Fuzzy-trace theory's analysis of recall, the optimization model, makes different predictions, owing to its emphasis on output interference and its assumptions about the differential sensitivity of verbatim and gist traces to such interference. In the first place, it is assumed, naturally, that dissociated verbatim and gist traces of targets are stored. It is further assumed (Brainerd, Howe, & Reyna 1994) that verbatim traces are necessary preconditions for free recall; that there is some probability of recalling a target if a trace of its physical presentation can be accessed, or if both verbatim and gist traces can be accessed, but not if only a gist trace can be accessed. Across trials, recall of individual targets tends to be based initially on verbatim readout and subsequently on reconstruction from gist (Brainerd et al. 1994). This brings us to the concept of *output interference,* which applies to verbatim traces, and to the concept of *memory strength,* which applies to gist.

Output interference is a response-produced process that varies across free-recall sequences as a function of the numbers and types of targets that have been recalled (e.g., Brainerd & Reyna 1989; Cooney & Troyer in press; Reyna & Brainerd 1989). Recalling an item generates off-task noise that impedes the retrieval of subsequent targets. This is output interference. Further, various findings suggest that accumulated output interference is more likely to impede the retrieval of verbatim traces than the retrieval of gist traces (Brainerd & Reyna 1993b; Reyna in press). So, other things being equal, it becomes more difficult to retrieve any verbatim trace as output interference accumulates. Moreover, differences in the retrieval difficulty of individual verbatim traces depend on their relative sensitivity to output interference. Hence, more interference-sensitive verbatim traces (higher error rates) will be more likely to be read out when accumulated interference is low (e.g., at the start of a test), and less interference-sensitive ones (lower error rates) will tend to be read out when accumulated interference is higher. This means that when verbatim-readout is operative, the ideal ordering should be harder (more interference sensitive) to easier (less interference sensitive) because output interference accumulates as recall proceeds.

Turning to memory strength, this is a stable property of the gist that subjects access in long-term memory on the basis of the inputs in recall tasks (Brainerd, Reyna, Howe, & Kevershan 1991). Various conceptualizations of this property can be found in the literature, such as echo intensity (Hintzman 1986), node interconnectedness (Anderson & Bower 1973), image strength (Shiffrin et al. 1990), and the like. For our purposes, however, the relevant points that fall out of these conceptualizations are (a) that the sequence in which gist traces are accessed during recall is stronger - - -> weaker and (b) that gist traces are relatively insensitive to accumulated output interference. It follows from Point a that when reconstruction from gist is the operative form of recall, the readout order will be easier - - -> harder. Whereas we saw that the verbatim-readout variety of recall will tend to be operative early in a free-recall protocol (when accumulated output interference is low), it follows from Point b that the gist-reconstruction form of recall can be operative later (when accumulated output interference is high). Thus, when the information that has been stored about targets consists of mixtures of verbatim traces for some and gist traces for others, recall of the former will tend to be concentrated at the start of output.

This analysis has three important implications. First, retrieval difficulty is not a unitary construct. It has a verbatim component and a gist component. Second, retrieval difficulty means interference sensitivity for the verbatim component and memory strength for the gist component. Third, because output interference accumulates as recall proceeds, the relationship between error rate (a composite measure of retrieval difficulty) and recall order should not be the easier - - -> harder sequence forecast by most theories. But, what should this relationship look like, and how should it change with age?

Recall Order: Data. The optimization model's answer is to predict five fine-grained effects: hard priority, overall nonmonotonicity, intertrial sharpening, developmental sharpening, and facilitation by difficulty grouping. The first two effects, in particular, constitute the operational definition of cognitive triage. The definitions of all the effects, along with the model's rationales for them, are as follows:

1. Hard priority. Suppose that the targets that are recalled on Trial 2 of a free-recall experiment are dichotomized as to retrieval difficulty using Trial 1 performance: "hard" = error on Trial 1 and "easy" = success on Trial 1. If the mean output positions of the two types of items are computed, the mean position ought to be earlier for hard items. The reasons are that (a) verbatim-based recall is concentrated at the start of output and that (b) targets for which only verbatim traces are accessible are more likely to produce an error on the preceding trial than those for which gist traces are also accessible (Brainerd et al. 1994). The same rationale applies to later pairs of adjacent trials.

2. Overall nonmonotonicity. Suppose that the targets that are recalled on some Trial i are classified as to relative retrieval difficulty using their complete error-success history from Trials 1 to $i - 1$. That is, easiest targets = 0 prior errors, next-easiest targets = 1 prior error, ..., hardest targets = $i - 1$ prior errors. If this graded measure of difficulty is plotted against targets' order of recall, the relationship should be harder - - -> easier - - -> harder. The reasons are that (a) output tends to begin with verbatim-based recall, proceeding in a harder - - -> easier direction, and that (b) output eventually switches to gist-based recall, proceeding in an easier - - -> harder direction.

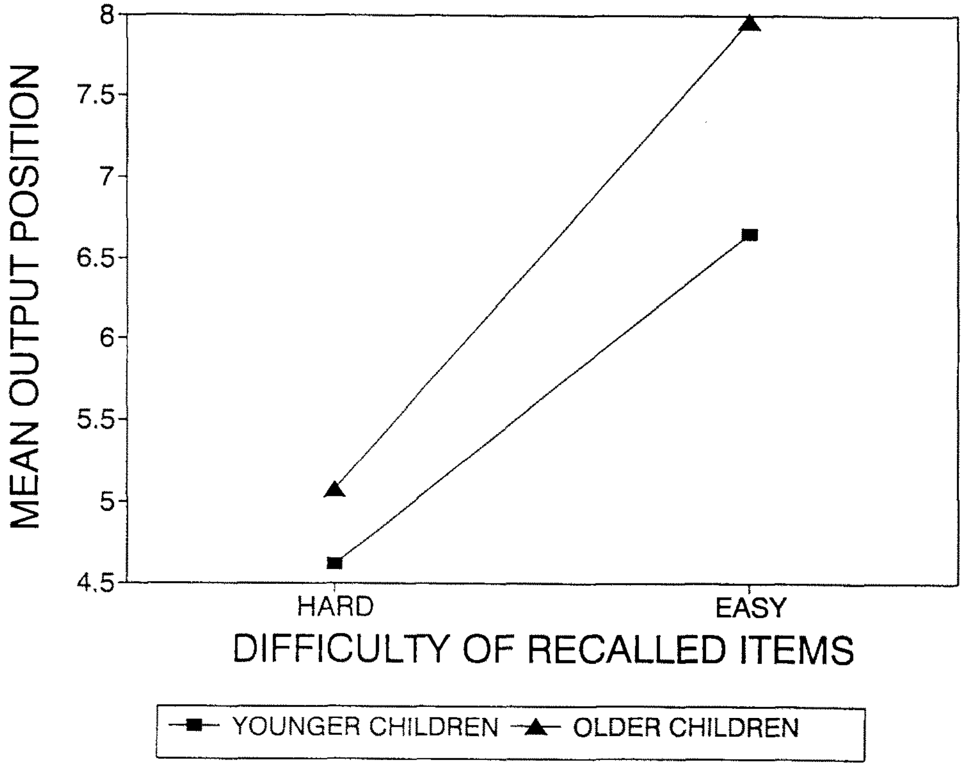
3. Intertrial sharpening. The hard priority and overall nonmonotonicity effects should become more pronounced as trials go by. The reasons are that we know empirically that recall improves across trials, and the optimization model says that both effects reflect processes that maximize total recall.

4. Developmental sharpening. The hard priority and overall nonmonotonicity effects should become more pronounced with age. Again, we known that recall improves with age, and the optimization model says that both effects reflect processes that maximize total recall.

5. Facilitation by difficulty clustering. Hard priority and overall nonmonotonicity are both effects in which subjects group their output in terms of the retrieval difficulty of targets; targets are recalled in bursts of items having similar difficulty levels. Free-recall protocols can be scored for the degree of such difficulty clustering. If hard priority and overall nonmonotonicity maximize output, as the optimization model claims, these difficulty-clustering scores should correlate with total recall.

We have examined these, and other, predictions in several developmental experiments concerned with the relationship between retrieval difficulty and recall order (Brainerd, Olney, & Reyna 1993; Brainerd, Reyna, Harnishfeger, & Howe 1993; Brainerd, Reyna, & Howe 1990; Brainerd, Reyna, Howe, & Kevershan 1990, 1991; Harnishfeger & Brainerd 1994). The hard priority and overall nonmonotonicity effects were initially isolated in experiments by Brainerd, Reyna, and Howe (1990) and by Brainerd, Reyna, Howe, and Kevershan (1990). The former effect is illustrated in Figure 6 and the latter in Figure 7 using data from some of those experiments. In Figure 6, recalled targets have been dichotomized as hard or easy, and their respective mean output positions have been plotted. As predicted, hard targets were articulated earlier in free recall than easy targets. In Figure 7, recalled targets

FIGURE 6
The hard priority effect in free recall at two age levels.

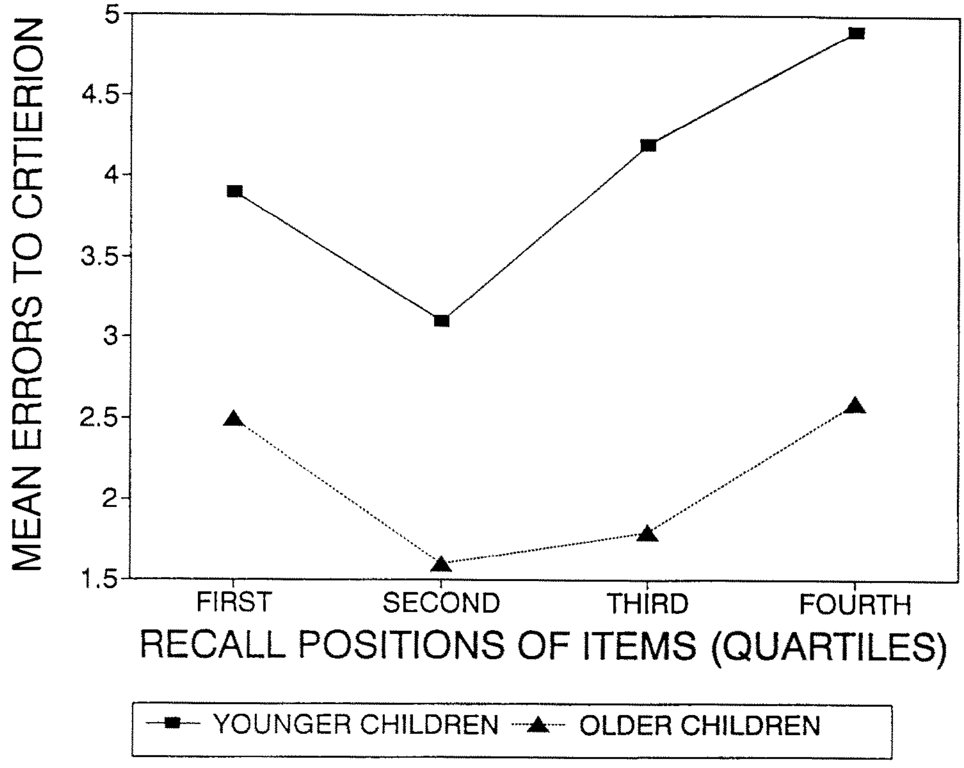


have been classified as to retrieval difficulty using a graded measure of difficulty (mean errors to criterion). As predicted, the hardest targets occupied the earliest and latest output positions, whereas the easiest targets occupied intermediate output positions. The third effect, intertrial sharpening, refers to the fact that both of the effects shown in Figures 6 and 7 tend to be less pronounced when they are based on data from earlier trials of an experiment.

Developmental sharpening is also illustrated in Figure 6 and Figure 7. The experiments on which these figures are based included subjects of two age levels, 7-year-olds and 11-year-olds. The data of each age level have been plotted separately, and it can be seen that the disparity in the output positions of targets with different difficulty classifications is greater in the older children.

The fifth effect, facilitation by difficulty grouping, has been studied in some of our most recent experiments. This effect is concerned with the question of whether cognitive triage, in the sense of phenomena such as hard priority and overall nonmonotonicity, is good for recall. The optimization model says that it should be because interference-sensitive verbatim traces are read out early, and interference-resistant gist traces are read out later. To test the claim, however, a goodness-of-tri-

FIGURE 7
The overall nonmonotonicity effect in free recall at two age levels.



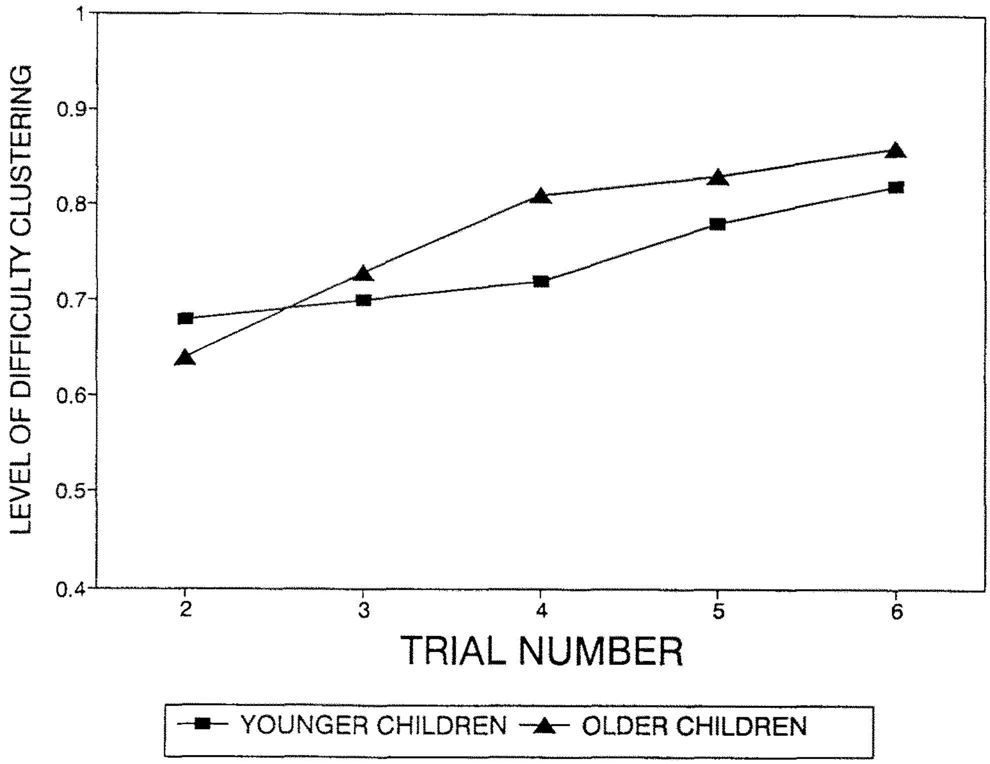
age measure of some sort is required that can be correlated with total recall scores. Brainerd, Reyna, Harnishfeger, and Howe (1993) developed a simple measure based on familiar category clustering statistics (e.g., Frankel & Cole 1971). They pointed out, first, that such statistics can be used whenever targets can be classified

TABLE 1
Correlations between Difficulty-Clustering Scores and Total Number of Targets Recalled in Brainerd, Reyna, Harnishfeger, and Howe (1993)

	Age		
Experiment	*Grade Two*	*Grade Six*	*College*
1	.35*	.57*	
2	.30*	.44*	
3			.49*
4	.28*	.41*	
5			.56*

*$p < .05$

FIGURE 8
Intertrial improvements in difficulty clustering in free recall at two age levels.



according to a few mutually-exclusive categories and that, second, retrieval difficulty is just such a classification scheme.

Brainerd, Reyna, Harnishfeger, and Howe (1993) reported five experiments, with subjects ranging from grade two to college age, in which clustering statistics were used to measure the tendency to group targets by difficulty level during free recall—specifically, to recall targets in clusters of hard and easy items. In all experiments and at all age levels, difficulty-clustering scores predicted accuracy; the higher the scores, the better the recall. Correlations are displayed by age level and experiment in Table 1. Note that in addition to being significant, the magnitude of these correlations increases with age. This is what the optimization model would predict if age improvements in recall accuracy are at least partly due to age improvements in cognitive triage. Another important finding is shown in Figure 8, where mean difficulty-clustering scores are plotted by trial and by age level. In line with the principles of intertrial and developmental sharpening, the optimization

TABLE 2
Some Alternative Explanations of Cognitive Triage

Model	*Definition*
Continuation	The output sequence on Trial 1 is easier – – –> harder. The output sequence on Trial 2 and later trials continues on with the hard end of the Trial 2 sequence, and eventually recycles to the easier – – –> harder part of the Trial 1 sequence. [See Brainerd (in press).]
Metamnemonic	Conceptually, subjects understand that some targets will be easier to retrieve than others. After Trial 1, they hold some of the hardest ones in short-term memory, read them out in a burst at the start of the next test, then recycle to easier - - -> harder for the rest of the test. [See Brainerd, Reyna, Howe, & Kevershan (1990).]
Difficulty Reordering	Following each trial, subjects devote special mnemonic processing to the hardest targets. This transforms some of them into the easiest targets. [See Brainerd, Olney, & Reyna (1993) and Brainerd, Reyna, & Howe (1990).]
Serial Position	Hard priority and overall nonmonotonicity are both artifacts of the recency effect in free recall. [See Brainerd, Reyna, Howe, & Kevershan (1991).]

model would naturally expect that these scores should increase across trials and with age. It can be seen that the data were consistent with both predictions.

In addition to testing the optimization models' fine-grained predictions about retrieval difficulty-output order phenomena, a key objective of cognitive triage experiments has been to test alternative models that might explain particular effects. It would be far beyond the scope of this paper to discuss those models and related findings in detail. We shall therefore confine the discussion to only a single alternative, the strategic model, that represents a fundamentally different outlook on recall. However, examples of other models are provided in Table 2, along with articles where relevant experiments may be found.

By "the strategic model," we simply mean the hypothesis that has dominated developmental and adult research on recall for the past quarter-century. That hypothesis is that developmental and individual differences in recall accuracy are primarily under the control of higher cognitive processes. The use of *organizational strategies*, either at input or at output, is the cognitive process that is emphasized in most accounts. The standard measures of this process are subjective-organization scores, which are calculated when subjects recall lists of unrelated targets, and category-clustering scores, which are calculated when subjects recall categorized lists. Consistent with the strategic model, it has long been known, at least in adults, that both types of scores correlate with total recall (Bjorklund & Muir 1988).

The optimization model is clearly a basic-processes theory rather than a high-cognitive theory. It assumes that recall is facilitated by the unconscious orchestration of sensitivity to output interference and memory strength. The model also provides a measure of the efficiency of this orchestration in the form of difficulty-clustering scores. We have seen that these scores, like subjective-organization scores and category-clustering scores, predict total recall. Brainerd, Reyna, Harnishfeger, and Howe (1993) observed that this fact allows one to pit the optimization model directly against the strategic model within individual data sets. When subjects of different ages study lists of unrelated targets, difficulty-clustering and subjective-organization scores can both be calculated for each free-recall protocol. The question then becomes, Which type of score shows more age variability and does a better job of predicting accuracy? When subjects of different ages study lists of targets that belong to a few familiar categories (e.g., animals, clothing), difficulty-clustering and category-clustering scores can both be calculated for each protocol. Again, the question is, which type of score shows more age variability and does a better job of predicting accuracy?

The first question was investigated by Brainerd, Reyna, Harnishfeger, and Howe (1993), and the second was investigated by Harnishfeger and Brainerd (1994). In the former case, we already know (Table 1) that (a) difficulty-clustering scores correlate with accuracy at all age levels, that (b) these correlations increase with age, and that (c) difficulty-clustering scores increase with age. Since all of these experiments involved free-recall of unrelated lists, subjective-organization scores were also calculated. There were three major findings, each of which favored the optimization model over the strategic model. First, subjective-organization scores failed to predict accuracy among either younger or older children. Second, although these scores predicted accuracy in college students, the relationship was weaker than for difficulty clustering. Third, among children, subjective-organization scores were not significantly greater than zero and did not increase with age. Fourth, although these scores were greater than zero in college students, they were much smaller than the corresponding difficulty-clustering scores.

Turning to the second question, Harnishfeger and Brainerd (1994) reported two categorized-recall experiments in which difficulty-clustering scores and category-clustering scores could both be computed. The subjects were between the ages of 7 and 13. As with unrelated lists, they found that difficulty-clustering scores correlated with accuracy in younger and older children, that the correlations increased with age, and that the magnitude of the scores increased with age. The picture was different for category-clustering scores. In contrast to subjective-organization scores, category-clustering scores were significantly greater than zero at all age levels, and they exhibited developmental improvement. However, category-clustering scores correlated with accuracy in only the oldest children. Also, they were smaller than difficulty-clustering scores at all age levels.

SUMMARY

As of the early 1980s, the relationship between higher reasoning and memory for critical informational inputs to reasoning was assumed to be captured by one or both of two hypotheses: memory necessity and constructivism. According to necessity, which was favored both by Piagetians and by information-processing theorists, accurate reasoning (e.g., solving a probability-judgment problem) demanded that faithful memories of critical problem information (e.g., the exact numerical values of sets) be retained at least long enough to complete the required sequence of reasoning operations. According to constructivism, which originated with Bartlett (1932) but was also favored by Piaget (e.g., Piaget & Inhelder 1973), the accuracy of memory for critical problem information was subordinated to the reasoner's efforts to make sense of that information.

Although necessity and constructivism were authorized by then-dominant theories of cognitive development, and also seemed to square with common-sense, they were repeatedly disconfirmed in experiments that began to appear in the mid-1980s. Those disconfirmations were of two general sorts. In one line of experimentation, when memory for critical background facts on childhood reasoning problems (e.g., the premises in a transitive inference problem, the cardinal numbers in a class-inclusion problem) was tested, it was usually found to be nonpredictive of the accuracy of solutions to those same problems. Even when the accuracy of children's memory and reasoning *were* correlated on particular problems, children were apparently not processing those memories in the course of reasoning. In the other line of experimentation, it was found that the content of memories that were actually being processed was quite different than the exact form of the inputs. On problems that seemed to call for executing precise computations on inputs' exact forms, reasoning was often based on crude patterns and distinctions that were instantiated by the inputs (i.e., their gist).

The original version of fuzzy-trace theory was an attempt to explain such counterintuitive findings. That version consisted of seven assumptions about the nature of reasoning, the nature of memories that are processed during reasoning, and relationships (especially developmental relationships) between the two domains. Those assumptions were: (a) *gist extraction* (children and adults mine the surface forms of inputs for senses and patterns); (b) *fuzzy-to-verbatim continua* (children's and adult's memories contain an assortment of reasoning-relevant information that varies in specificity, ranging from traces that preserve the exact form of inputs to traces that preserve only vague distinctions); (c) *fuzzy-processing preference* (reasoning usually prefers to operate on memories that are near the fuzzy ends of fuzzy-to-verbatim continua); (d) *reconstruction in short-term memory* (immediate memory performance, as well as performance after a forgetting interval, can be based on reconstruction from stored gist); (e) *output interference* (sensitivity to the competing information that is generated on memory and reasoning tasks is a source of errors, particularly of children's errors, on such tasks); (f) *resource freedom* (that reasoning did not draw on a limited pool of generic resources) and (g) *ontogenesis* (factors a–e vary systematically with age).

Some of these assumptions have subsequently been refined in light of new data. For example, Assumptions a and c were refined via the identification of a range of relationships between memory and reasoning. Independence, as well as positive and negative dependency, have now been brought under experimental control; each of these relationships can be produced by varying the relative accessibility of verbatim versus gist representations. As a second example, Assumption b was clarified by the discovery that subjects systematicaly adjusted the level of precision of their reasoning in different tasks. As a third example, Assumptions c and g were refined via the identification of specific gists that formed the basis of solutions to particular problems and of developmental patterns in reliance on those gists. As a fourth example, Assumptions e and g were clarified by the discovery of competing gists in certain tasks and of developmental declines in sensitivity to such interference. As a final example, Assumptions a and b were clarified by the conclusion that in order to explain available data on memory-reasoning relationships, it was necessary to assume that gist memories are not extracted from verbatim memories of inputs but, rather, the two are stored in parallel.

Four areas in which research on fuzzy-trace theory is currently concentrated are suggestibility and false memories, judgment and decision making, the development of forgetting, and the development of retrieval. In the first area, phenomena such as inferential recall, false recognition, reversals of false recognition, say/mean confusions, and the suggestibility of memory through internal (autosuggestion) and external influences (the presentation of misleading information) have been studied. A general conclusion that has emerged from this work is that some of the phenomena that have been treated as different are actually similar (i.e., inferential recall, false recognition, say/mean confusions, some source confusions, and autosuggestion), whereas other phenomena treated as similar are actually different (some source confusions and various suggestibility effects). Regarding the latter, effects due to forgetting of original verbatim memories, to forgetting of verbatim details about source, and to retrieval of gist in lieu of verbatim memories are distinguished. Although suggestibility and false memory encompass multiple phenomena, they can all be explained by the distinction between verbatim and gist memories of the same inputs, and the different ways in which those memories behave.

In the second area, judgment and decision making have been explored. Differences between judgment and decision making are predicted by fuzzy-trace theory's assumptions regarding task calibration (i.e., that differences depend on the required specificity of the response). Research on judgment, in particular, has indicated that the complexities of inclusion relationships can produce a chain of errors in representation, retrieval, and processing, providing a unitary account of class-inclusion reasoning, conjunction fallacies, base-rate neglect, conditional probability judgment, deductive reasoning, and other tasks. Research on decision making has contrasted predictions of traditional theories, which claim that numerical processing is central to observing decision phenomena, with those of fuzzy-trace theory. We found that the processing of such numerical information is neither necessary nor sufficient to produce standard decision phenomena. Decision biases were attributed to the propensity for intuitive reasoning, which was shown to increase with age. This has led to reconceptualizations of two old questions, the nature of cognitive competence and the nature of rationality. Concerning competence, rather than characterizing competence as either present or absent, fuzzy-trace theory identifies different levels of understanding. In order of sophistication, these range from failing to know a reasoning principle, to failure to appreciate the relevance of that principle in varied contexts, to failure to recognize and correct misdirections in reasoning that are due to interference. Concerning rationality, gist memories are assumed to be the key to rationality because they allow reasoning to be invariant across situations that differ in superficial details.

In the third area, age variability in rates of forgetting from long-term memory, factors that might mask age variability, factors that might produce artifactual age variability, and possible theoretical sources of age variability have been investigated. The major theoretical conclusions so far are that storage failures contribute more to forgetting rates than retrieval failures do, that this is true throughout the life span, and that loss of verbatim memories contribute more to storage failures than loss of gist memories do, and this is also true throughout the life span.

In the fourth area, we have investigated the sequence in which stored information is accessed in long-term memory during unconstrained retrieval and some factors that affect the access sequence. This research has produced the surprising findings that weak, rather than strong, memory representations tend to be accessed first and that this tendency becomes more pronounced as development proceeds. The explanation seems to lie in the fact that recall generates proactive interference, in the form of off-task noise associated with accessing and articulating individual items, and that verbatim memories of items are more sensitive to such interference than gist memories. The latter point again underscores fuzzy-trace theory's claim that the distinction between verbatim and gist memories of the same inputs, and the various assumptions that are associated with this distinction, are fundamental to understanding a broad range of memory and reasoning phenomena.

ACKNOWLEDGMENTS: Much of the research that is reviewed in this article was supported by grants from the Natural Sciences and Engineering Research Council and the Spencer Foundation.

REFERENCES

Ackerman, B. P. (1992). "The sources of children's source errors in judging causal inferences." *Journal of Experimental Child Psychology, 54*, 90–119.

Acredolo, C., J. O'Connor, L. Banks, & K. Horobin. (1989). "Children's ability to make probability estimates: Skills revealed through application of Anderson's functional measurement methodology." *Child Development, 60*, 933–945.

Alba, J. & L. Hasher. (1983). "Is memory schematic?" *Psychological Bulletin, 93*, 203–231.

Anderson, J. R. & G. H. Bower. (1973). *Human associative memory*. Washington, DC: Winston.

Argyris, C. (1988). Problems in producing usable knowledge for implementing liberating alternatives. Pp. 540–561 in *Decision making: Descriptive, normative, and prescriptive inter-*

actions, edited by D. E. Bell, H. Raiffa, & A. Tversky. New York: Cambridge University Press.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

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

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.

Beach, L. R. (1990). *Image theory: Decision making in personal and organizational contexts*. New York: Wiley.

Beal, C. R. (1990a). "Development of knowledge about the role of inference in text comprehension." *Child Development, 61*, 1011–1023.

———. (1990b). "The development of text evaluation and revision skills." *Child Development, 61*, 247–258.

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

Belli, R. F., D. S. Lindsay, M. S. Gales, & T. T. McCarthy. (1994). "Memory impairment and source misattribution in postevent misinformation experiments with short retention intervals." *Memory & Cognition, 22*, 40–54.

Bernoulli, D. (1738/1954). "Specimen theoraie novae de mensura sortis." *Commentarii Academiae Sientiarum Imperiales Petropolitanae, 5*, 175–192.

Binet, A. (1900). *La suggestibilite*. Paris: Schleicher Freres.

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

———. (1989). *Children's thinking: Developmental and individual differences*. Pacific Grove, CA: Brooks/Cole.

Bjorklund, D. F. & K. K. Harnishfeger. (1987). "Developmental differences in mental effort requirements for the use of an organizational strategy in free recall." *Journal of Experimental Child Psychology, 44*, 109–125.

———. (1990). "The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition." *Developmental Review, 10*, 48–71.

Bjorklund, D. F. & J. E. Muir. (1988). "Children's development of free recall memory: Remembering on their own." *Annals of Child Development, 5*, 79–123.

Brainerd, C. J. (1981). "Working memory and the developmental analysis of probability judgment." *Psychological Review, 88*, 463–502.

———. (1983). "Young children's mental arithmetic errors: A working-memory analysis." *Child Development, 54*, 812–830.

———. (1985a). "Model-based approaches to storage and retrieval development." Pp. 143–208 in *Basic processes in memory development*, edited by C. J. Brainerd & M. Pressley. New York: Springer-Verlag.

———. (1985b). "Three-state models of memory development: A review of advances in statistical methodology." *Journal of Experimental Child Psychology, 40*, 375–394.

———. (1986, November). Development of forgetting from long-term memory. Paper presented at Psychonomic Society, Boston.

———. (in press). "Children's forgetting, with implications for memory suggestibility." In *Memory for everyday and emotional events*, edited by N. L. Stein, P. A. Ornstein, B. Tversky, & C. J. Brainerd. Hillsdale, NJ: Erlbaum.

———. (1995). "Interference processes in memory development: The case of cognitive triage." In *Interference and inhibition in cognition*, edited by F. N. Dempster & C. J. Brainerd. San Diego, CA: Academic Press.

Brainerd, C. J. & L. L. Gordon. (1994). "Development of verbatim and gist memory for numbers." *Developmental Psychology, 30*, 163–177.

Brainerd, C. J., M. L. Howe, & V. F. Reyna. (1994). Fuzzy-trace theory and the development of recall. Manuscript submitted for publication.

Brainerd, C. J. & J. Kingma. (1984). "Do children have to remember to reason? A fuzzy-trace theory of transitivity development." *Developmental Review, 4*, 311–377.

———. (1985). "On the independence of short-term memory and working memory in cognitive development." *Cognitive Psychology, 17*, 210–247.

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

Brainerd, C. J., C. A. Olney, & V. F. Reyna. (1993). "Optimization versus effortful processing in children's cognitive triage: Criticisms, reanalyses, and new data." *Journal of Experimental Child Psychology, 55*, 353–373.

Brainerd, C. J. & P. A. Ornstein. (1991). "Children's memory for witnessed events: The developmental backdrop." In *The suggestibility of children's recollections*, edited by J. Doris. Washington, DC: American Psychological Association.

Brainerd, C. J. & V. F. Reyna. (1988a). "Generic resources, reconstructive processing, and children's mental arithmetic." *Developmental Psychology, 24*, 324–334.

———. (1988b). "Memory loci of suggestibility development: Comment on Ceci, Ross, and Toglia." *Journal of Experimental Psychology: General, 117*, 197–200.

———. (1989). "Output-interference theory of dual-task deficits in memory development." *Journal of Experimental Child Psychology, 47*, 1–18.

———. (1990a). "Can age x learnability interactions explain the development of forgetting?" *Developmental Psychology, 26*, 194–204.

———. (1990b). "Gist is the grist: Fuzzy-trace theory and the new intuitionism." *Developmental Review, 10*, 3–47.

———. (1990c). "Inclusion illusions: Fuzzy-trace theory and perceptual salience effects in cognitive development." *Developmental Review, 10*, 365–403.

———. (1991). "Acquisition and forgetting processes in normal and learning-disabled children: A disintegration/redintegration theory." Pp. 147–178 in *Advances in the neuropsychology of learning disabilities*, edited by J. E. Obrzut & G. W. Hynd. San Diego, CA: Academic Press.

———. (1992a). "Explaining memory-free reasoning." *Psychological Science, 3*, 332–339.

———. (1992b). "The memory independence effect: What do the data show? What do the theories claim?" *Developmental Review, 12*, 164–186.

———. (1993a). "Domains of fuzzy-trace theory." (Vol. 1), Pp. 50–94 in *Emerging themes in cognitive development*, edited by M. L. Howe & R. Pasnak. New York: Springer-Verlag.

———. (1993b). "Memory independence and memory interference in cognitive development." *Psychological Review, 100*, 42–67.

———. (in press a). "Autosuggestibility in memory development." *Cognitive Psychology.*

———. (in press b). "Learning rate, learning opportunities, and the development of forgetting." *Developmental Psychology.*

Brainerd, C. J., V. F. Reyna, K. K. Harnishfeger, & M. L. Howe. (1993). "Is retrieval grouping good for recall?" *Journal of Experimental Psychology: General, 122*, 249–268.

Brainerd, C. J., V. F. Reyna, & M. L. Howe. (1990). "Cognitive triage in children's memories: Optimal retrieval or effortful processing?" *Journal of Experimental Child Psychology, 49,* 428–447.

Brainerd, C. J., V. F. Reyna, M. L. Howe, & J. Kevershan. (1990). "The last shall be first: How memory strength affects children's retrieval." *Psychological Science, 1,* 247–252.

———. (1991). "Fuzzy-trace theory and cognitive triage in memory development." *Developmental Psychology, 27,* 351–369.

Brainerd, C. J., V. F. Reyna, M. L. Howe, & J. Kingma. (1990). "The development of forgetting and reminiscence." *Monographs of the Society for Research in Child Development, 53,* 3–4 (Whole No. 222).

Brainerd, C. J., V. F. Reyna, & R. Kneer. (in press). "False-recognition reversal: When similarity is distinctive." *Journal of Memory and Language.*

Bransford, J. D. & J. J. Franks. (1971). "The abstraction of linguistic ideas." *Cognitive Psychology, 2,* 331–380.

Bransford, J. D., J. R. Barclay, & J. J. Franks. (1972). "Sentence memory: A constructive versus interpretative approach." *Cognitive Psychology, 3,* 193–209.

Brouwer, L. E. J. (1952). "Historical background, principles and methods of intuitionism." *South African Journal of Science, 49,* 139–146.

Brown, A. L., S. S. Smiley, J. D. Day, M. A. R. Townsend, & S. C. Lawton. (1977). "Intrusion of a thematic idea in children's comprehension and retention of stories." *Child Development, 48,* 1454–1466.

Bryant, & T. Trabasso. (1971). "Transitive inferences and memory in young children." *Nature, 232,* 457–459.

Callahan, P. (1989). Learning and development of probability concepts: Effects of computer assisted instruction and diagnosis. Unpublished doctoral dissertation, University of Arizona, College of Education, Tucson.

Carey, S. & R. Diamond. (1977). "From piecemeal to configurational representations of faces." *Science, 195,* 312–341.

Case, R. (1985). *Intellectual development.* New York: Academic Press.

Cassel, W. S. & D. F. Bjorklund. (1994). Developmental patterns of eyewitness memory and suggestibility: An ecologically-based short-term longitudinal study. Manuscript submitted for publication.

Ceci, S. J. & M. Bruck. (1993). "Suggestibility of the child witness: A historical review and synthesis." *Psychological Bulletin, 113,* 403–439.

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

Chandler, C. C. (1991). "How memory for an event is influenced by related events: Interference in modified recognition tests." *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 115–125.

———. (in press). "Accessing related events increases retroactive interference in a matching recognition test." *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Chapman, M. & U. Lindenberger. (1992). "Transitivity judgments, memory for premises, and models of children's reasoning." *Developmental Review, 12,* 124–163.

Clark, H. H. & E. V. Clark. (1977). *Psychology and language.* New York: Harcourt Brace Jovanovich.

Cohen, J. (1981). "Can human irrationality be experimentally demonstrated?" *The Behavioral and Brain Sciences, 4,* 317–331.

Coombs, C. H. (1975). "Portfolio theory and the measurement of risk." In *Human judgment and decision processes*, edited by M. Kaplan & S. Schwartz. New York: Academic Press.

Cooney, J. B. & H. L. Swanson. (1990). "Individual differences in memory for mathematical story problems: Memory span and problem perception." *Journal of Educational Psychology, 82*, 570–577.

Cooney, J. B. & R. Troyer. (in press). "A dynamic model of reaction time in a short-term memory task." *Journal of Experimental Child Psychology.*

Daneman, M. & P. A. Carpenter. (1980). "Individual differences in working memory and reading." *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Dempster, F. N. (1985). "Short-term memory development in childhood and adolescence." In *Basic processes in memory development*, edited by C. J. Brainerd & M. Pressley. New York: Springer-Verlag.

———. (1992). "The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging." *Developmental Review, 12*, 45–75.

Dempster, F. N. & C. J. Brainerd. (Eds.). (1995). *Interference and inhibition in cognition.* San Diego, CA: Academic Press.

Estes, W. K. (1980). "Is human memory obsolete?" *American Scientist, 68*, 62–69.

Fagley, N. S. & P. M. Miller. (1987). "The effects of decision framing on choice of risky vs. certain options." *Organizational Behavior and Human Decision Processes, 39*, 264–277.

Fajnsztejn-Pollack, G. (1973). "A developmental study of decay rate in long-term memory." *Journal of Experimental Child Psychology, 16*, 225–235.

Falmagne, R. J. (Ed.). (1975). *Reasoning: Representation and process in children and adults.* Hillsdale, NJ: Erlbaum.

Fischer, G. W. & S. A. Hawkins. (1993). "Strategy compatibility, scale compatibility, and the prominence effect." *Journal of Experimental Psychology: Human Perception and Performance, 19*, 580–597.

Fisher, R. P. & C. C. Chandler. (1991). "Independence between recalling interevent relations and specific events." *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 722–733.

Flagg, P. W. (1976). "Semantic integration in sentence memory?" *Journal of Verbal Learning and Verbal Behavior, 15*, 491–504.

Fletcher, C. R. (1992). "Assessing recognition memory for surface forms in discourse: A methodological note." *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 199–203.

Flexer, A. J. & E. Tulving. (1978). "Retrieval independence in recognition and recall." *Psychological Review, 85*, 153–172.

Foley, M. A., M. K. Johnson, & C. L. Raye. (1983). "Age-related confusion between memories for thoughts and memories for speech." *Child Development, 54*, 51–60.

Frankel, F. & M. Cole. (1971). "Measures of category clustering in free recall." *Psychological Bulletin, 76*, 39–44.

Fulginiti, J. V. & V. F. Reyna. (1993, November). Qualitative versus quantitative processing in decision making: A fuzzy-trace theory analysis. Paper presented at the annual meeting of the Society for Judgment and Decision Making, Washington, DC.

Glucksberg, S. & J. Danks. (1975). *Experimental psycholinguistics: An introduction.* Hillsdale, NJ: Erlbaum.

Goldstein, W. M. & H. Einhorn. (1987). "Expression theory and the preference reversal phenomenon." *Psychological Review, 94*, 236–254.

Guttentag, R. E. (1989). "Age differences in dual-task performance: Procedures, assumptions, and results." *Developmental Review, 9,* 146–170.

Harlow, H. F. (1949). "The formation of learning sets." *Psychological Review, 56,* 51–65.

Harnishfeger, K. K. & C. J. Brainerd. (1994). "Nonstrategic facilitation of children's recall: Evidence of triage with semantically related information." *Journal of Experimental Child Psychology, 57,* 259–280.

Hasher, L. & R. T. Zacks. (1988). "Working memory, comprehension, and aging: A review and a new view." Pp. 193–224 in *The psychology of learning and motivation* (Vol. 22), edited by G. H. Bower. New York: Academic Press.

Henle, M. (1962). "On the relation between logic and thinking." *Psychological Review, 69,* 366–378.

Heyting, H. (1959). *Constructivity in mathematics.* Amsterdam: North-Holland.

Hilbert, D. (1923). "Die logischen grundlagen der mathematik." *Mathematische Annalen, 88,* 151–165.

Hintzman, D. L. (1986). "Schema abstraction in a multiple-trace memory model." *Psychological Review, 93,* 411–428.

Hoemann, H. W. & B. M. Ross. (1982). "Children's understanding of probability concepts." In *Children's logical and mathematical cognition,* edited by C. J. Brainerd. New York: Springer-Verlag.

Hogarth, R. M. (1980). *Judgment and choice: The psychology of decision.* Chicester, England: Wiley.

Howe, M. L. (1991). "Misleading children's story recall: Forgetting and reminiscence of the facts." *Developmental Psychology, 27,* 746–762.

Howe, M. L. & C. J. Brainerd. (1989). "Development of long-term retention." *Developmental Review, 9,* 301–340.

Howe, M. L., M. L. Courage, & L. Bryant-Brown. (1993). "Reinstating preschoolers' memories." *Developmental Psychology, 29,* 854–869.

Howe, M. L., M. L. Courage, & C. Peterson. (in press). "How can I remember when I wasn't there: Long-term retention of traumatic experiences and emergence of the cognitive self." *Cognition and Consciousness.*

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

Howe, M. L., A. Kelland, L. Bryant-Brown, & S. L. Clark. (1992). "Measuring the development of children's amnesia and hypermnesia." Pp.56–102 in *Development of long-term retention,* edited by M. L. Howe, C. J. Brainerd, & V. F. Reyna. New York: Springer-Verlag.

Howe, M. L., F. M. Rabinowitz, & M. J. Grant. (1993). "On measuring (in)dependence of cognitive processes." *Psychological Review, 100,* 737–747.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century-Crofts.

Jacobs, J. E. & M. Potenza. (1991). "The use of judgment heuristics to make social and object decisions: A developmental perspective." *Child Development, 62,* 166–178.

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

Kahneman, D., P. Slovic, & A. Tversky. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kahneman, D. & A. Tversky. (1979). "Prospect theory: An analysis of decision under risk." *Econometrica, 47,* 263–291.

Kail, R. (1993). "The role of global mechanisms in developmental change in speed of processing." Pp. 97–119 in *Emerging themes in cognitive development: 1. Foundations* edited by M. L. Howe & R. Pasnak. New York: Springer-Verlag.

Kendler, H. H. & T. S. Kendler. (1962). "Vertical and horizontal processes in problem solving." *Psychological Review, 69*, 1–16.

Keren, G. & W. A. Wagenaar. (1987). "Violation of utility theory in unique and repeated gambles." *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 387–391.

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

Kintsch, W., D. Welsch, F. Schmalhofer, & S. Zimny. (1990). "Sentence memory: A theoretical analysis." *Journal of Memory and Language, 29*, 133–159.

Klayman, J. (1985). "Children's decision strategies and their adaptation to task characteristics." *Organizational Behavior and Human Performance, 35*, 179–201.

Kolodner, J. L. (1983). "Reconstructive memory: A computer model." *Cognitive Science, 7*, 281–328.

Kreindler, D. & C. Lumsden. (in press). "Extracting a narrative's causal gist: A modeling study based on fuzzy-trace theory." *Journal of Experimental Child Psychology.*

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

Liben, L. S. & C. J. Posnansky. (1977). "Inferences on inference: The effects of age, transitive ability, memory load, and lexical factors." *Child Development, 1977, 48*, 1490–1497.

Lindsay, D. S. (1990). "Misleading suggestions can impair eyewitnesses' ability to remember event details." *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1077–1083.

Lindsay, D. S. & M. K. Johnson. (1991). "Recognition memory and source monitoring." *Bulletin of the Psychonomic Society, 29*, 203–205.

Lim, P. (1993). Meaning versus verbatim memory in language processing: Deriving inferential, morphological, and metaphorical gist. Unpublished doctoral dissertation, University of Arizona.

Loftus, E. F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.

Loftus, E. F., B. Levidow, & S. Duensing. (1992). "Who remembers best? Individual differences in memory for events that occurred in a science museum." *Applied Cognitive Psychology, 6*, 93–107.

Loftus, E. F. & G. R. Loftus. (1980). "On the permanence of stored information in the human brain." *American Psychologist, 35*, 409–420.

Lopes, L. L. (1987). "Between hope and fear: The psychology of risk." *Advances in Experimental Social Psychology, 20*, 255–295.

Lovett, S. B. & J. A. Singer. (1991, April). The development of children's understanding of probability: Perceptual and quantitative conceptions. Paper presented at Society for Research in Child Development, Seattle, WA.

Marbe, K. (1901). *Experimentell-psychologische untersuchungen über das utriel.* Leipzig: Engelmann.

Marche, T. A. (1993). On the resistance of preschoolers' memories to postevent misinformation. Unpublished doctoral dissertation, Memorial University.

McGeoch, J. A. (1932). "Forgetting and the law of disuse." *Psychological Review, 39*, 352–370.

———. (1942). *The psychology of human learning.* New York: Longmans, Green.

McNeil, B. J., S. G. Pauker, S. G. Sox, & A. Tversky. (1982). "On the elicitation of preferences for alternative therapies." *New England Journal of Medicine, 306,* 1259–1262.

Merriman, W. E., M. Azmita, & M. Perlmutter. (1988). "Rate of forgetting in early childhood." *International Journal of Behavioral Development, 11,* 467–474.

Miller, G. A. (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information." *Psychological Review, 63,* 81–97.

Moore, C. F., J. A. Dixon, & B. A. Haines. (1991). "Components of understanding in proportional reasoning: A fuzzy set representation of developmental progressions." *Child Development, 62,* 441–459.

Morrison, F. J., M. M. Haith, & J. Kagan. (1980). "Age trends in recognition memory for pictures: The effects of delay and testing procedure." *Bulletin of the Psychonomic Society, 16,* 480–483.

Navon, D. (1984). "Resources—a theoretical soup stone?" *Psychological Review, 91,* 216–234.

Nisbett, R. E. & L. Ross. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Offenbach, S. I., G. E. Gruen, & B. J. Caskey. (1984). "Development of proportional response strategies." *Child Development, 55,* 963–972.

Ornstein, P. A. & K. Corsale. (1979). "Organizational factors in children's memory." In *Organization, structure, and memory,* edited by C. R. Puff. New York: Academic Press.

Paris, S. G. & A. Y. Carter. (1973). "Semantic and constructive aspects of sentence memory in children." *Developmental Psychology, 9,* 109–113.

Paris, S. G. & B. K. Lindauer. (1977). "Constructive aspects of children's comprehension and memory." Pp. 35–60 in *Perspectives on the development of memory and cognition,* edited by R. V. Kail & J. W. Hagen. Hillsdale, NJ: Erlbaum.

Payne, J., J. R. Bettman, & E. J. Johnson. (1988). "Adaptive strategy selection in decision making." *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 534–552.

———. (1992). "Behavioral decision research: A constructive processing perspective." *Annual Review of Psychology, 43,* 87–131.

Penfield, W. & L. Roberts. (1959). *Speech and brain mechanisms.* Princeton, NJ: Princeton University Press.

Pennington, N. & R. Hastie. (1993). "The story model for juror decision making." Pp. 192–221 in *Inside the juror: The psychology of juror decision making,* edited by R. Hastie. New York: Cambridge University Press.

Perner, J. & D. G. Mansbridge. (1983). "Developmental differences in encoding length series." *Child Development, 54,* 710–719.

Piaget, J. (1949). *Traite de logique.* Paris: Colin.

———. (1953). *Logic and psychology.* Manchester, England: Manchester University Press.

———. (1970). *Genetic epistemology.* New York: Columbia University Press.

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

Poole, D. A. & D. S. Lindsay. (in press). "Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and learning questions on reports of nonexperienced events." *Journal of Experimental Child Psychology.*

Poole, D. A. & L. T. White. (1991). "Effects of question repetition on the eyewitness testimony of children and adults." *Developmental Psychology, 27,* 975–986.

———. (1993). "Two years later: Effects of question repetition and retention interval on the eyewitness testimony of children and adults." *Developmental Psychology, 29,* 844–853.

———. (in press). "Tell me again and again: Stability and change in the repeated testimonies of children and adults." In *Memory, suggestibility, and eyewitness testimony in children and adults*, edited by M. Zaragoza, J. R. Graham, G. N. N. Hall, R. Hirschman, & Y. S. Ben-Porath. Thousand Oaks, CA: Sage.

Prawatt, R. S. & A. Cancelli. (1976). "Constructive memory in conserving and nonconserving first graders." *Developmental Psychology, 12*, 47–50.

Rabinowtiz, F. M., M. L. Howe, & J. A. Lawrence. (1989). "Class inclusion and working memory." *Journal of Experimental Child Psychology, 48*, 379–409.

Reder, L. M. (1982). "Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification." *Psychological Review, 89*, 250–280.

Reitman, J. S. & G. H. Bower. (1973). "Storage and later recognition of exemplars of concepts." *Cognitive Psychology, 4*, 194–206.

Reyna, V. F. (1981). "The animated word: Modification of meaning by context." *Dissertation Abstracts International, 42*, 3852B [Ann Arbor, MI: 82–03153].

———. (1985). "Figure and fantasy in children's language." In *Cognitive learning and memory in children: Progress in cognitive development research*, edited by M. Pressley & C. Brainerd. New York: Springer-Verlag.

———. (1990a, November). Categorical versus compensatory reasoning and inconsistent choice. Paper presented at the 31st Annual Meeting of the Psychonomic Society, New Orleans, LA.

———. (1990b, June). Fuzzy phrases and fuzzy choices. Invited presentation at the Third Annual Nags Head Conference on Judgment and Decision Making, Kill Devil Hills, NC.

———. (1991). "Class inclusion, the conjunction fallacy, and other cognitive illusions." *Developmental Review, 11*, 317–336.

———. (1992a). "Reasoning, remembering, and their relationship: Social, cognitive, and developmental issues." Pp. 103–127 in *Development of long-term retention* edited by M. L. Howe, C. J. Brainerd, & V. F. Reyna. New York: Springer-Verlag.

———. (1992b, November). Selective processing effects and theories of choice. Paper presented at the 33rd Annual Meeting of the Psychonomic Society. St. Louis, MO.

———. (1995). "Interference effects in memory and reasoning: A fuzzy-trace theory analysis." In *Interference and inhibition in cognition*, edited by F. N. Dempster & C. J. Brainerd. San Diego, CA: Academic Press.

———. (in press). "Meaning, memory, and the interpretation of metaphors." In *Metaphor: Pragmatics and applications*, edited by J. Mio & A. Katz. Hillsdale, NJ: Erlbaum.

Reyna, V. F. & C. J. Brainerd. (1989a). Fuzzy-trace theory of framing effects in choice. Presented at the 30th Annual Meeting of the Psychonomic Society, Atlanta, GA.

———. (1989b). "Output interference, generic resources, and cognitive development." *Journal of Experimental Child Psychology, 47*, 42–46.

———. (1990). "Fuzzy processing in transitivity development." *Annals of Operations Research, 23*, 37–63.

———. (1991a). "Fuzzy-trace theory and children's acquisition of mathematical and scientific concepts." *Learning and Individual Differences, 3*, 27–59.

———. (1991b). "Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversion." *Journal of Behavioral Decision Making, 4*, 249–262.

———. (1992). "A fuzzy-trace theory of reasoning and remembering: Paradoxes, patterns, and parallelism." Pp. 235–259 in *From learning processes to cognitive processes: Essays in* *honor of William K. Estes* (Vol. 2), edited by A. Healy, S. Kosslyn, & R. Shiffrin. Hillsdale, NJ: Erlbaum.

———. (1993). "Fuzzy memory and mathematics in the classroom." In *Everyday memory,* edited by R. Logie & G. Davies. Amsterdam: North Holland.

———. (in press). "The origins of probability judgment: A review of data and theories." In *Subjective Probability,* edited by G. Wright & P. Ayton. New York: Wiley.

Reyna, V. F. & S. C. Ellis. (in press). "Fuzzy-trace theory and framing effects in children's risky decision making." *Psychological Science.*

Reyna, V. F. & J. Fulginiti. (1992). Ambiguity, precision, and choice: A fuzzy-trace analysis. Paper presented at the Annual Meeting of the Society for Judgment and Decision Making. St. Louis, MO.

Reyna, V. F. & B. Kiernan. (1994a). Children's memory and interpretation of psychological metaphors. Manuscript in preparation.

———. (1994b). "The development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval." *Developmental Psychology, 30,* 178–191.

Russell, B. (1903). *The principles of mathematics.* Cambridge, England: Cambridge University Press.

Salthouse, T. A. (1990). "Working memory as a processing resource in cognitive aging." *Developmental Review, 10,* 101–124.

Schmidt, C. R. & M. Welch. (1989, March). The development of rote and gist memory tendencies in young children. Paper presented at Society for Research in Child Development, Kansas City, MO.

Schneider, W. & M. Pressley. (1989). *Memory development from 2 to 20.* New York: Springer-Verlag.

Schwartz, B. & D. Reisberg. (1991). *Learning and memory.* New York: W. W. Norton.

Shiffrin, R. M., R. Ratcliff, & S. E. Clark. (1990). "List-strength: II. Theoretical mechanisms." *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 179–195.

Shafir, E. B., D. N. Osherson, & E. E. Smith. (1989). "An advantage model of choice." *Journal of Behavioral Decision Making, 2,* 1–23.

Shipley, E. F. (1979). "The class-inclusion task: Question form and distributive comparisons." *Journal of Psycholinguistic Research, 8,* 301–331.

Siegler, R. S. (1981). "Developmental sequences within and between concepts." *Monographs of the Society for Research in Child Development, 46,* (Whole No. 189).

———. (1988). "Strategy choice procedures and the development of multiplication skill." *Journal of Experimental Psychology: General, 117,* 258–275.

———. (1991). *Children's thinking.* Englewood Cliffs, NJ: Prentice-Hall.

Simon, H. A. (1956). "Rational choice and the structure of the environment." *Psychological Review, 63,* 129–138.

———. (1975). "The functional equivalence of problem solving skills." *Cognitive Psychology, 7,* 268–288.

Small, M. Y. & J. Butterworth. (1981). "Semantic integration and the development of memory for logical inferences." *Child Development, 52,* 732–735.

Smedslund, J. (1969). "Psychological diagnostics." *Psychological Bulletin, 71,* 237–248.

Spinillo, A. G. & P. Bryant. (1991). "Children's proportional judgments: The importance of half." *Child Development, 62,* 427–440.

Surber, C. F. & B. A. Haines. (1987). "The growth of proportional reasoning." *Annals of Child Development, 4*, 35–87.

Titcomb, A. L. & V. F. Reyna. (1995). "Memory interference and misinformation effects." In *Interference and inhibition in cognition*, edited by F. N. Dempster & C. J. Brainerd. San Diego, CA: Academic Press.

Trabasso, T. (1977). "The role of memory as a system in making transitive inferences." In *Perspectives on the development of memory and cognition*, edited by R. V. Kail, Jr., & J. W. Hagen. Hillsdale, NJ: Erlbaum.

Turner, M. L. & R. W. Engle. (1989). "Is working memory capacity task dependent?" *Journal of Memory and Language, 28*, 127–154.

Turvey, M. T. (1974). "Constructive theory, perceptual systems, and tacit knowledge." Pp. 165–180 in *Cognition and the symbolic processes*, edited by W. B. Weimer & D. S. Palermo. Hillsdale, NJ: Erlbaum.

Tversky, A. (1969). "Intransitivity of preferences." *Psychological Review, 76*, 31–48.

———. (1972). "Elimination by aspects: A theory of choice." *Psychological Review, 79*, 281–299.

Tverksy, A. & D. Kahneman. (1981). "The framing of decisions and the psychology of choice." *Science, 211*, 453–458.

———. (1983). "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment." *Psychological Review, 91*, 293–215.

———. (1986). "Rational choice and the framing of decisions." *Journal of Business, 59*, 252–278.

———. (1992). "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty, 5*, 297–323.

Tversky, A., S. Sattath, & P. Slovic. (1988). "Contingent weighting in judgment and choice." *Psychological Review, 95*, 371–384.

Underwood, B. J. (1954). "Speed of learning and amount retained: A consideration of methodology." *Psychological Bulletin, 51*, 276–282.

———. (1964). "Degree of learning and the measurement of forgetting." *Journal of Verbal Learning and Verbal Behavior, 3*, 112–129.

Van Wallendael, L. R. & R. Hastie. (1990). "Tracing the footsteps of Sherlock Holmes: Cognitive representations of hypothesis testing." *Memory & Cognition, 18*, 240–250.

von Neumann, J. & O. Morgenstern. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wagner, S., E. Winner, D. Cicchetti, & H. Gardner. (1981). "Metaphorical mapping in human infants." *Child Development, 52*, 728–731.

Warren, A. R. & P. L. Lane. (in press). "Effects of timing and type of questioning on eyewitness accuracy and suggestibility." In *Memory, suggestibility, and eyewitness testimony in children and adults*, edited by M. Zaragoza, J. R. Graham, G. N. N. Hall, R. S. Hirschman, & Y. S. Ben-Porath. New York: Harper & Row.

Wedell, D. H. & U. Bockenholt. (1990). "Moderation of preference reversals in the long run." *Journal of Experimental Psychology: Human Perception and Performance, 16*, 429–438.

———. (1992). The effects of multiple plays on preference reversals: An examination of stimulus, task, and subject variables. Manuscript submitted for publication.

———. (in press). "Contemplating single versus multiple encounters of a risky prospect." *American Journal of Psychology*.

Werner, H. (1948). *Comparative psychology of mental development*. New York: International Universities Press.

Wilkinson, A. C. (1982). "Theoretical and methodological analysis of partial knowledge." *Developmental Review, 2,* 274–304.
Winer, G. A. (1980). "Class-inclusion reasoning in children: A review of the empirical literature." *Child Development, 51,* 309–328.
Winer, G. A. & C. McGlone. (1993). "On the uncertainty of conservation: Reponses to misleading conservation questions." *Developmental Psychology, 29,* 760–769.
Wolfe, C. R. (1994). Information seeking on Bayesian conditional probability problems: A fuzzy-trace theory account. Manuscript submitted for publication.