

## Developmental Reversals in False Memory: A Review of Data and Theory

C. J. Brainerd, V. F. Reyna, and S. J. Ceci  
Cornell University

Can susceptibility to false memory and suggestion increase dramatically with age? The authors review the theoretical and empirical literatures on this counterintuitive possibility. Until recently, the well-documented pattern was that susceptibility to memory distortion had been found to decline between early childhood and young adulthood. That pattern is the centerpiece of much expert testimony in legal cases involving child witnesses and victims. During the past 5 years, however, several experiments have been published that test fuzzy-trace theory's prediction that some of the most powerful forms of false memory in adults will be greatly attenuated in children. Those experiments show that in some common domains of experience, in which false memories are rooted in meaning connections among events, age increases in false memory are the rule and are sometimes accompanied by net declines in the accuracy of memory. As these experiments are strongly theory-driven, they have established that developmental improvements in the formation of meaning connections are necessary and sufficient to produce age increases in false memory.

*Keywords:* false memory, fuzzy-trace theory, child witnesses, opponent processes

False memories are quintessential episodic memory phenomena: One remembers specific events as having happened during some passage of life, when in fact the events did not happen then, if ever. False memories are often harmless—as when one remembers serving a bottle of Merlot with dinner when actually it was Zinfandel. There are circumstances, however, in which such errors are far from benign. Consider, for instance, an emergency room physician who prescribes treatments based on false memories of patients' symptoms (e.g., that a patient reported shortness of breath), a not uncommon occurrence (Reyna & Lloyd, 1997). It is in the legal arena, however, where one encounters perhaps the most widely acknowledged source of harmful false memories. In the courtroom, the bulk of the evidence that bears on guilt or innocence is presented through sworn testimony, which means that it is true only to the extent that a witness's recollections are accurate. Although it is sometimes possible to disprove errant recollections with unimpeachable forensic evidence (e.g., fingerprints, DNA results), that is rare because forensic evidence is gathered in a remarkably small percentage of crimes—less than 10% in the United States, by some estimates (e.g., Horvath & Meesig, 1996, 1998). Not surprisingly, studies of known cases of false conviction have revealed that some prototypical forms of legal false memory, such as false identifications of innocent suspects (e.g., Wells, Small, Penrod, Malpass, Fulero, & Brimacombe, 1998) and false recollections during interrogations (e.g.,

Kassin, 2005), are leading causes of false convictions (Brainerd & Reyna, 2005).

Since the 1980s, much experimentation has revolved around another source of legal false memories—namely, false memories of child witnesses. The forensic implications of children's false memories have been topics of keen interest since the beginning of scientific psychology (e.g., Binet, 1900; Small, 1896; Stern, 1910; Varendonck, 1911). Nearly a century ago, Whipple published a series of influential reviews of this early work in *Psychological Bulletin* (Whipple, 1909, 1911, 1912, 1913), concluding that young children's memories are so susceptible to distortion as to render their testimony prejudicial, even when recalling events of great personal moment. The impetus for the new work that began in the 1980s was the fact that, in the immediately preceding years, long-standing legal barriers to child witnesses had fallen, and evidence from children was becoming increasingly frequent in certain types of cases (for a review, see McGough, 1993). The Federal Rules of Evidence 601 (The Committee on the Judiciary of the House of Representatives, 2006), adopted by most jurisdictions in criminal cases, broadened the admissibility of children's testimony, allowing the trier of fact to decide how much weight to give to it (rather than excluding it statutorily). The most common examples of child-related cases are domestic crimes in which children are either habitual victims (e.g., abuse and neglect) or habitual witnesses (e.g., dependency hearings, custody disputes, spousal violence, home production of controlled substances). Had it not been for another historical development, however, admissibility of evidence from children probably would not have stimulated intense research on their false memories. That development was a series of high-profile cases in which defendants were tried for multiple counts of graphic and often bizarre sexual abuse on the basis of child allegations that raised reliability questions in the scientific as well as the legal community (see Ceci & Bruck, 1995, for descriptions of such cases).

---

C. J. Brainerd, V. F. Reyna, and S. J. Ceci, Department of Human Development, Cornell University.

Preparation of this article was supported in part by National Science Foundation Grant SBR973-0143 and National Institutes of Health Grant R01-MH021211. We are indebted to G. S. Goodman and M. P. Toglia for their helpful comments on a draft of this article.

Correspondence concerning this article should be addressed to C. J. Brainerd, Department of Human Development, Cornell University, Ithaca, NY 14853. E-mail: cb299@cornell.edu

*State of New Jersey v. Michaels* (1994) is a prominent illustration. The defendant, a 26-year-old preschool teacher, was convicted on 115 counts of sexual abuse involving 20 children and sentenced to 47 years in prison. Many of the children's allegations were implausible, including claims that the defendant played the piano while nude, made children eat her feces, and raped them with knives, forks, and Lego blocks. Pursuant to a review of this case, a committee of 46 concerned scientists concluded that the children's reports of abuse may have been tainted by false memories that had been stimulated by suggestive interviewing procedures. The conviction was reversed following the committee's submission of an *amicus* brief in support of an appeal to the New Jersey Supreme Court (Bruck & Ceci, 1995), but not until the defendant had served 4 years in prison.<sup>1</sup>

An impressive research literature on children's false memories has now accumulated, the results of which have been periodically reviewed and integrated (e.g., Bruck & Ceci, 1999; Ceci & Bruck, 1993, 1995; Goodman, 2006; Goodman & Schaaf, 1997; Holliday, Reyna, & Hayes, 2002; Quas, Qin, Schaaf, & Goodman, 1997; Reyna, Mills, Estrada, & Brainerd, 2007). That literature contains many findings of theoretical and forensic significance. Forensically, for example, some of the results have been exploited to devise model investigative interviewing protocols that maximize the yield of true information from children while minimizing the yield of false information (for reviews, see Hershkowitz, Fisher, Lamb, & Horowitz, 2007; Lyon & Saywitz, 2006; Poole & Lamb, 1998). There is general consensus, though, that the findings of greatest overall importance are those that establish global age trends in false memory. Until recently, the picture that had emerged was quite consistent. False memories had been found to decrease with age throughout childhood and adolescence (Chae & Ceci, 2005; McFarlane, Powell, & Dudgeon, 2002; Young, Powell, & Dudgeon, 2003), and age was the best single predictor of individual variability in false memories (Ceci, Papierno, & Kulkofsky, 2007). Such age declines have been reported in any number of memory suggestion experiments whose designs were intended to emulate the manipulative interviewing practices that first fomented scientific interest in children's false memories (e.g., Ackil & Zaragoza, 1995, 1998; Bjorklund, Bjorklund, & Brown, 1998; Bjorklund et al., 2000; Eisen, Qin, Goodman, & Davis, 2002; Goodman, Quas, Batterman-Faunce, Riddlesberger, & Kuhn, 1994; Holliday, & Hayes, 2000, 2001; Marche, 1999; Marche & Howe, 1995). The same pattern has been detected with other false-memory paradigms that do not employ suggestion. Examples include developmental studies of sentence recognition (e.g., Reyna & Kiernan, 1994, 1995), memory for narratives (e.g., Ackerman, 1992, 1994), word recognition (e.g., Brainerd & Reyna, 1996; Brainerd, Reyna, & Kneer, 1995), free and cued recall of live event sequences (e.g., Pipe, Gee, Wilson, & Egerton, 1999; Poole & White, 1991), free and cued recall of word lists (e.g., Bjorklund & Muir, 1988), and memory for mathematical propositions (Brainerd & Reyna, 1995).

Age declines in false memory have direct theoretical and forensic ramifications. On the theoretical side, this pattern supports a traditional view in which age improvements in accuracy are the touchstone of memory development, with early childhood being seen as a time when the boundary between reality and fantasy is indistinct and useful mnemonic strategies have not yet been acquired (Brainerd, Reyna, & Forrest, 2002). As Goodman (2006)

noted, "there are a host of theoretical issues to be addressed. Additional research that tests theoretical explanations for why young children are more suggestible and have less accurate memories than adults is still a priority" (p. 830). The modal conception is that cognitive factors (e.g., source-monitoring), social factors (e.g., susceptibility to persuasion), metacognitive factors (e.g., introspective awareness of memory states), and neurobiological factors (e.g., maturation of the prefrontal cortex) all contribute. The underlying logic is that because each of these factors promotes memory accuracy and resistance to suggestion, they converge to ensure age declines in false memory. Below, however, we show that there are systematic exceptions to the general claim of age declines in false memory and that, more important, these exceptions are predicted by an established theory of false memory.

On the forensic side, scientific documentation of age declines in false memory has figured prominently in many court cases in which the reliability of young children's recollections and their competence to provide evidence were central issues (e.g., Bruck & Ceci, 1995; Ceci & Friedman, 2000; Lyon, 1995). In such cases, children's competence to testify has been contested in pre-trial motions and hearings, and their courtroom testimony has been challenged as being unreliable, all on the ground that their statements are far more likely than adults' to be infected with false memories (Brainerd & Reyna, 2005). To take a concrete illustration, consider a common occurrence in certain types of cases: There are two versions of events that bear on a defendant's guilt, an exculpatory version and an incriminating version, with one being provided by adults and the other being provided by children. The jury's obligation in this circumstance is to render a verdict by evaluating the relative credibility of the different versions of events. If children's reports are more likely than adults' to contain false memories, it follows that, other factors being equal, the adult version ought to be assigned a higher credibility rating. This is precisely the basis for much expert testimony. Consider the following example of such testimony:

Well, in virtually all these studies, two and three-year olds do not do well in suggestibility, and the four and five-year olds . . . do pretty well.

It's true that the sorts of questioning that were asked of the children are not supported by basic research into suggestibility, but these children were all over the age of 6, the cut-off for suggestibility-proneness in scientific studies. (Expert testimony by a prosecution witness *In the Matter of Riley, Shelby, and Austin Blanchard v. John Blanchard*, 2001; Tr. pp. 1,441)

Because age declines in false memory and age increases in net accuracy are so thoroughly documented, it may seem odd to ask whether there are domains of experience that display marked age *increases* in false memory, perhaps coupled with net *declines* in accuracy. However, one theory that is used to explain adults' false memories, fuzzy-trace theory (FTT), predicts just such counterintuitive results for domains of experience that are ubiquitous in real life (Brainerd & Reyna, 1998a; Ceci & Bruck, 1998). Moreover, many studies have recently appeared that confirm such predictions

<sup>1</sup> Although the New Jersey Supreme Court justices did not cite the *amicus* brief explicitly, they quoted heavily from it in their opinion to reverse the defendant's convictions.

and put the theoretical mechanisms from which they sprang to the test. Like the ubiquitous age-decrease pattern, a contrasting age-increase pattern would have broad implications. Most obviously, it would sharply constrain the modal theoretical conception of memory development and the aforementioned forensic interpretations. Theoretically, it would be necessary to explain how, in some domains of experience, errors could decline and overall accuracy could improve while, in other domains, the opposite is true. Forensically, the default principle that children's evidence is more likely than adults' to be tainted by false memories would not be tenable. In its place, a more refined view that takes additional findings about memory development into account would be necessary.

Since the initial studies of developmental reversals in false memory were reported, there has been rapid progress on both the empirical and theoretical fronts. The aim of this article is to review what has been learned so far. In *Review of Theory: Opponent Processes and False-Memory Development*, we sketch the theoretical distinctions that first led researchers to expect developmental reversals in false memory, and we show how such reversals fall out of those distinctions. In *Review of Data: Three Lines of Evidence*, we review findings from experiments in which predicted developmental reversals have been verified, with a view toward extracting global patterns. Data from three types of designs are considered: (a) developmental studies of false memory for Deese/Roediger/McDermott (DRM) materials (Deese, 1959; Roediger & McDermott, 1995); (b) developmental studies of false memory for categorized materials; and (c) developmental studies of false memory following suggestive questioning. In *Testing Process Models of Developmental Reversals*, further results are reviewed that supply tests of the principles that predict developmental reversals in false memory. Those principles consist of specific mechanisms that stimulate errant recollections along with other mechanisms that suppress them. We emphasize findings from manipulations that are able to establish causal relations between these mechanisms and developmental reversals in false memory. Data that bear on relations of necessity and sufficiency are both considered. Finally, in *Concluding Remarks: The Next Steps*, we examine some of the ramifications of this review for basic and forensic research on children's false memories and describe a series of targets for future research.

Before proceeding to the main business of this article, we offer an advance observation on the merits of theory-guided research on applied questions—in this instance, the reliability of children's memory reports. Long ago, Kurt Lewin advised psychologists that scientific progress on applied questions flows from the realization that “there is nothing so practical as a good theory” (1951, p. 169). The findings that we review in the upcoming sections illustrate that Lewin's advice is as probative today as it was when he first gave it. As will be seen, because studies of developmental reversals in false memory have been strongly theory-driven, the literature has advanced at a rapid pace. For instance, the fact that these counterintuitive age trends were not accidental discoveries but, rather, were deliberately sought on theoretical grounds has meant that a relatively modest amount of experimentation has been able to establish their reality. Further, because this search is simultaneously a test of theoretical predictions, it has been natural to include experimental manipulations that embody the mechanisms that generated those predictions in the first place. The denouement

is a literature that has not only produced solid evidence of counterintuitive age trends but has also advanced developmental theory and forensic practice by establishing causal connections between those trends and the memory processes that are ostensibly responsible for them.

### Review of Theory: Opponent Processes and False-Memory Development

We consider three topics in this section. First, in order to establish a common methodological language, the generic features of false-memory experiments are described. Second, the core principles of FTT's account of false memory are summarized. Last, we show how those principles have been used to create a recipe for detecting developmental reversals in false memory.

#### *How False Memory Is Measured*

False-memory experiments measure episodic memories, albeit erroneous ones: The subject's task is to report information that was encountered in a specific exposure context, with principal interest attaching to whether they also report unrepresented information that resembles presented information in certain respects (which are specified by experimental design). In the canonical example of such an experiment, subjects are exposed to a collection of memory targets (words, pictures, sentences, narratives, video images, live event sequences), and then they respond to recognition or recall tests. For instance, suppose that the targets are objects and events in a story about a child who attended a birthday party. If recall tests are administered, children either report everything they can remember about the target material (free recall) or report portions of the material in response to specific cues (e.g., “tell me what the children ate,” “tell me what the children wore on their heads,” “tell me what the children drank”).

If recognition tests are administered, test probes are of three basic types: (a) targets (e.g., eating chocolate ice cream, putting on a blue birthday hat, drinking lemonade); (b) unrepresented items that preserve salient features of targets (most often, meaning features: eating strawberry ice cream, putting on a red birthday hat, drinking punch); and (c) further unrepresented items that are nominally unrelated to the targets (e.g., petting a dog, answering the telephone, playing a video game). With the exception of occasional studies using nonsense materials (e.g., Brainerd, Stein, & Reyna, 1998), the events that these probes tap are all familiar ones, which is to say that they are well-defined items that are represented in children's long-term memories. The last two types of probes are usually called related distractors and unrelated distractors, respectively. Type b probes are the false-memory items because, although they were not presented, a part of their content was presented (ice cream, but not strawberry; birthday hat, but not red; drink, but not punch). None of the content of the unrelated distractors was presented, and, hence, these items provide baseline measures of response bias.

If recall tests are administered, the index of false memory is simply the intrusion rate, which is the proportion of recalled information that was not presented as part of the designated exposure episode (e.g., Poole & White, 1991; Toglia, Neuschatz, & Goodwin, 1999). Although intrusions usually preserve the meaning of target materials (Brainerd & Reyna, 2005; Bjorklund

& Muir, 1988), they sometimes preserve other features of targets, such as their phonology or orthography. In some studies, therefore, intrusions are decomposed into errors that preserve distinct features of target materials (e.g., Dewhurst & Robinson, 2004). If recognition tests are administered, the measurement of false memory is less direct because recognition responses can be based on guessing and other forms of response bias, rather than memory. Therefore, the index of false memory is not the false-alarm rate for related distractors, but rather it is the false-alarm rate for related distractors relative to the false-alarm rate for unrelated distractors. Because each probe must be classified as old (targets) or new (related and unrelated distractors), various sources of response bias contribute to the hit rate and false-alarm rate (e.g., Snodgrass & Corwin, 1988). Logically, significant levels of false memory can be said to be present only if the false-alarm rate for related distractors (false memory + response bias) exceeds the false-alarm rate for unrelated distractors (response bias only). This means that signal detection statistics of memory discrimination, such as  $A'$  or  $d'$ , are better measures of false memory than are raw false-alarm rates for related distractors. The use of such statistics is essential in developmental research because the influence of response bias declines between early childhood and young adulthood (e.g., Reyna & Kiernan, 1994). Owing to that datum, although raw false-alarm rates for related distractors inflate estimates of false memory at any age level, the inflation is more pronounced at younger than at older age levels. This, in turn, can mask age increases in false memory and can manufacture spurious age declines (for examples, see Brainerd & Mojardin, 1998).

In addition to age trends in false memory, age trends in net accuracy can also be measured. Of the total amount of information that is recalled or recognized (both true and false), net accuracy is the ratio of true information to true information plus false information. In most experimental paradigms, true recall and true recognition will increase with age, naturally. As long as false recall and false recognition decline with age, net accuracy must increase. However, if false recall and false recognition increase with age, net accuracy may decrease or increase or remain unchanged, accordingly as the age increase in false recall or recognition is greater than, less than, or the same as the age increase in true recall or recognition. We shall see, below, that some recent developmental studies of false memory have reported age declines in net accuracy. This is an instructive pattern because it narrows the range of viable theoretical interpretations. When studies show that false memory increases with age while net accuracy declines, it cannot be argued that the former finding is somehow due to older subjects' superior ability to remember all aspects of target materials. Rather, if false memory increases disproportionately, the age increases must be due to older subjects' superior ability to remember specific aspects of target materials that are likely to distort remembering (e.g., semantic relations, see below).

Summing up the discussion so far, the false memories that are measured in the canonical developmental study are memories of items that children have experienced many times in the past but that, by experimental design, they did not experience during the episodic context that is the focus of recognition or recall tests. There is one further methodological distinction that is crucial—namely, the distinction between studies that focus on spontaneous false memories versus studies that focus on implanted false memories (see Reyna & Lloyd, 1997). In studies of spontaneous false

memories, no attempt is made to foment false recognition or recall of related distractors. Children are simply exposed to the target material, and at some later point, they respond to recognition or recall tests. Therefore, observed levels of false memory are due to endogenous distortion processes (e.g., meaning-driven reconstruction), and age trends in false memory reflect age trends in those endogenous processes (Brainerd & Reyna, 1998a). Studies of implanted false memories incorporate an additional design feature: a misinformation phase that is interpolated between the exposure and test phases, whose aim is to implant false memories of specific related distractors. In an attempt to emulate the investigative practices that originally stimulated interest in children's false memories, the misinformation phase is usually disguised as an interview about (or a review of) objects and events from the focal episodic context (e.g., the birthday party story). However, the interview contains leading questions that are intended to implant false memories of particular related distractors (e.g., "Remember when the children ate the strawberry ice cream?" "Remember when the children put on the red birthday hats?" "Remember when the children drank the punch?"). Thus, observed levels of false memory for these items are due to some combination of the influences of endogenous distortion processes and exogenous distortion, and age trends in false memory reflect age trends in both mechanisms (Brainerd & Reyna, 1998a).

The developmental literature on false memory contains many studies of both the spontaneous and implanted variety, but the relative mix of the two designs has shifted over the years. A decade ago, misinformation studies dominated the data archive—the reason, again, being scientific concerns about the consequences of suggestive forensic interviews. In recent years, however, spontaneous designs have become more common, a shift that has been motivated by theoretical and forensic considerations (Brainerd & Reyna, 2005). Theoretically, a vital feature of any paradigm is its ability to establish a clear mapping between false-memory responses and the underlying processes that cause them. The mapping is more straightforward in spontaneous designs than in misinformation designs. False-memory responses in spontaneous designs are due to a single group of processes (endogenous distortion mechanisms), whereas in misinformation designs, the same responses confound two classes of processes (endogenous distortion mechanisms and suggestion-induced distortion). In misinformation designs, controls must therefore be incorporated if the effects of the two groups of processes are to be separated, which adds methodological complexity. Moreover, there is a long-standing question as to whether the false-memory responses in misinformation designs are due to actual memory distortion or to nonmemorial factors, especially susceptibility to social influence (Ceci & Bruck, 1993). Additional controls must be incorporated to take account of this latter possibility, which further complicates misinformation designs (Reyna & Brainerd, 1998). A key forensic motivation for developmental studies that implement the spontaneous procedure is that the scientific information that courts need to evaluate the reliability of children's evidence has changed as investigative interviewing practices have evolved. Research showing that children's memories are highly susceptible to misinformation prompted a national movement to reduce the incidence of suggestive questioning (McGough, 1993; see various chapters in Zaragoza, Graham, Hall, Hirschman, & Ben-Porath, 1995). Ultimately, as mentioned earlier, this led to the creation of the model



interviewing protocols that secure substantial amounts of information from even very young children, without resorting to suggestive questioning (e.g., Orbach & Lamb, 2000, 2001; Sternberg, Lamb, Esplin, Orbach, & Mitchell, 2001). Although suggestive child interviewing is a continuing problem in criminal investigation (e.g., *State of Arizona v. Gregory Speers*, 2007), as the use of nonsuggestive protocols has become more widespread, evaluating the reliability of children's evidence has come to require scientific information on another question: Are children's baseline tendencies to report false memories considerably higher than adolescents' or adults', even when they have not been suggestively questioned? Developmental studies of spontaneous false memory supply data on this question.

### *Opponent-Processes Principles*

While developmental research has been motivated by forensic questions, adult research has been primarily concerned with testing the predictions of theoretical models of false memory. As noted in recent reviews (Brainerd & Reyna, 2005; Reyna et al., 2007), although there are older theories of false memory—the three major ones being constructivism (Bransford & Franks, 1971), schema theory (Alba & Hasher, 1983), and the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993)—the dominant theories in the current literature are opponent-processes models. Such models evolved to explain several findings that older theories did not predict. Although many examples of such findings might be given (see Brainerd & Reyna, 2005; Reyna & Lloyd, 1997), a general class of results that violates these older theories is statistical and experimental dissociations between true and false memories. The older theories are often referred to as one-process models inasmuch as they posit that the same types of memory representations underlie both true and false memory, which leads to the prediction that the two will be positively correlated within individuals and will react similarly to experimental manipulations (Reyna & Lloyd, 1997). However, it has often been found that true and false memory are stochastically independent within individuals (e.g., Brainerd & Reyna, 1995; Brainerd et al., 1995; Reyna & Kiernan, 1994), and many manipulations have been identified that dissociate them (i.e., that affect one type of performance without affecting the other or that drive them in opposite directions). Several examples of dissociation manipulations figure in a later section of this article (see Testing Process Models for Developmental Reversals). In contrast to older theories, statistical and experimental dissociations are core predictions of opponent-processes models.

In this subsection, we briefly describe, first, the historical lineage of such models, followed by a brief summary of how a model that has been widely used in both adult and developmental research explains false alarms and intrusions of related distractors. Opponent-processes conceptions grew out of work on dual-trace models of episodic memory. In the latter models, episodic traces are viewed as collections of features, with a core distinction being drawn between surface and semantic features—as in, for example, the processing implicit and explicit representations (PIER) model of Nelson and associates (Nelson, Schreiber, & McEvoy, 1992), the retrieving efficiently from memory (REM) model of Shiffrin and associates (Shiffrin, 2003; Shiffrin & Steyvers, 1997), and FTT (Reyna & Brainerd, 1995).<sup>2</sup> Surface features are representations of identifying characteristics that are shared by exemplars of

targets (e.g., the shape, color, and size of an object), some of which are sampled when an exemplar is encountered (e.g., during the birthday party story). Semantic features are representations of the meaning qualities of targets (e.g., that it is a sweet dessert food), some of which are also sampled when an exemplar is encountered. Thus, when target materials are presented, individual targets can be thought of as picking out (identifying, selecting, activating) their representations in long-term memory, which results in the sampling of subsets of surface and semantic features from those representations. A further important notion is that sampled features become episodically tagged, that is, that features are labeled with contextual cues (e.g., visual, auditory, temporal, emotional) from the situations in which the exemplars are encountered. This is what it means, in these models, to store episodic traces of experience: Sampled features are tagged with contextual cues and stored, which provides records of what happened during that episode that can be retrieved on memory tests. A final important principle is that separate episodic traces of targets' surface and semantic features are stored. Examples of separate episodic traces of surface and semantic features are the verbatim and gist traces of FTT, the explicit and implicit traces of PIER, and the episodic and lexical-semantic traces of REM. Such distinctions have been exploited to account for dissociations among memory tasks that vary in their degree of emphasis on accessing the surface form of study materials versus accessing their meaning content (Nelson, McGivney, Gee, & Janczura, 1998; Reyna & Lloyd, 1997; Shiffrin, 2003).

In addition to these dual-trace assumptions, opponent-processes conceptions incorporate the assumption that the two types of episodic traces support opposite responses to related distractors (see Brainerd & Reyna, 2005). Similar to the presentation of targets during the exposure phase, the presentation of a related distractor on a recognition test or its mental generation on a recall test activates its corresponding representation in long-term memory, resulting in feature sampling for the related distractor. It also provokes retrieval of episodic traces of the earlier target presentations. Related distractors (e.g., *punch*) are items that share many semantic features (but few surface features) with certain targets (e.g., *lemonade*), and they share few surface or semantic features with other targets (e.g., *blue birthday hat*). Hence, whether or not subjects falsely recognize or recall a related distractor depends on what types of episodic traces are retrieved, and more explicitly, it depends on whether there is a high level of match between the features that are sampled for the related distractor and the features that are stored in retrieved episodic traces of targets (Brainerd & Reyna, 2005). On the one hand, suppose that a related distractor (*punch*) provokes retrieval of episodic traces of unrelated targets (*blue birthday hat*). The level of match will be low because there is little overlap between the features that are stored in these episodic traces and the features that are sampled for the related distractor. On the other hand, suppose that the same related distractor provokes retrieval of episodic traces of related targets (*lemonade*). Now, the level of match will depend on whether

<sup>2</sup> There are several important differences between REM, PIER, and FTT when it comes to their respective characterizations of the content of dual episodic traces. Those differences are not pertinent here, however, and the theories' assumptions about representation can be treated as equivalent for present purposes.

verbatim or gist traces are retrieved. (For simplicity, we will use the verbatim/gist distinction of FTT rather than the implicit/explicit distinction of PIER or the episodic/lexical–semantic distinction of REM, in the remainder of this article, though the three distinctions are equivalent for present purposes.) The level of match will be high, supporting false-memory responses, if semantic traces are retrieved because, by design, there will be considerable overlap between the semantic features that are sampled for related distractors and the semantic features that are stored in gist traces of related targets (e.g., *punch* and *sweet drink*, respectively). However, the level of match will be low, suppressing false-memory responses, if verbatim traces are retrieved. This is because, again by design, there will be salient differences in the surface features that are sampled for related distractors and the surface features that are stored in verbatim traces of related targets (e.g., the word *punch* bears no literal resemblance to the word *lemonade*). This process, wherein the high familiarity of related distractors is neutralized by verbatim traces of their corresponding targets, is called recollection rejection in FTT, and its properties have been discussed by Brainerd, Reyna, Wright, and Mojardin (2003). The point to which interest attaches here, however, is that retrieval of episodic traces of semantic features supports false memories while retrieval of episodic traces of surface features does not, which is responsible for the “opponent” in opponent processes.

In this conception, false-memory responses will be increased by experimental procedures that strengthen gist traces relative to verbatim traces, while simultaneously, false-memory responses will be decreased by procedures that strengthen verbatim traces relative to gist traces. In the adult literature, one of the most common methods of strengthening gist relative to verbatim traces is to present multiple exemplars of a certain type of target (e.g., *lemonade*, *orange juice*, *apple juice*, *pineapple juice*). This increases false memory for related distractors (*punch*) because, of course, the gist trace of the relevant meaning (“sweet drink”) is strengthened by each successive target, but a target’s verbatim trace is not because each target is different. In the remainder of this article, we will refer to such procedures as connected-meaning paradigms, to convey that the elevations in false memory show that adults are able to connect meaning across distinct exemplars. Connected-meaning paradigms are especially important in the study of false memory, for two reasons. First and most obviously, the formation of semantic relations between objects and events is prominent in most situations in everyday life, by which we mean that those situations contain multiple objects and events that converge on the same meaning. (For instance, reflect on the large numbers of exemplars of fruit, vegetables, meat, candy, soda, seasonings, bread, pie, ice cream, and cheese you can expect to encounter on your next trip to the grocery store.) In everyday life, therefore, intrusions and false alarms that are supported by meaning connection are garden variety false memories (Did you see pumpkin pie in the freezer case? Did the bakery have French bread?). Second, these garden variety false memories are predicted to increase with age as more and stronger connections are acquired among objects and events (e.g., *citrus*, *dairy*, *predators*, etc. are categories that young children do not yet apprehend).

### *A Recipe for Developmental Reversals*

If the expression of false memories turns on which types of episodic traces, verbatim or gist, are retrieved, the standard age-

decrease pattern cannot be a universal trend, unless developmental changes in verbatim and gist memory are constrained in specific ways. Explicitly, it cannot be a universal trend unless, on the one hand, verbatim memory improves with age, so that the memory evidence that suppresses errors becomes more accessible and, on the other hand, gist memory does not improve with age, so that the memory evidence that foments errors does not become more accessible. However, developmental studies of verbatim and gist memory provide ample evidence that both improve between early childhood and young adulthood and that improvements in the two are surprisingly independent of each other (see Bouwmeester, Vermunt, & Sijtsma, 2007; Brainerd & Reyna, 2004; Reyna, Holliday, & Marche, 2002).

If the bases for suppressing and accepting false memories both improve with age, the standard age-decrease trend must be task dependent (Brainerd & Reyna, 1998a; Ceci & Bruck, 1998). The general theoretical expectation is that whether errors decrease with age, increase with age, or remain constant will depend on the mix of verbatim and gist memories that are tapped by individual false-memory paradigms. The more specific theoretical predictions are, first, that the standard age-decline pattern is favored in paradigms that are more sensitive to age variability in verbatim memory than to age variability in gist memory and, second, that the reverse pattern is favored in paradigms that are more sensitive to age variability in gist memory than to age variability in verbatim memory. Because the age-decline pattern has been reported in so many studies, the key questions at this point are the following: Do the designs of these studies have the features that are specified in the first prediction? Are there any simple design modifications that would yield unambiguous tests of the second prediction? We consider these two questions separately.

### *Verbatim Sensitivity and Age Declines in False Memory*

The first question has been examined in recent articles (e.g., Brainerd & Reyna, 2007; Brainerd et al., 2002; Reyna et al., 2007). There, it has been shown that a ubiquitous feature of developmental studies reduces sensitivity to age improvements in gist memory but not sensitivity to improvements in verbatim memory: Great care is taken to present only target materials whose meanings are well understood, even by very young children. That is, although the exact target materials (e.g., word lists, sentence lists, narratives, sequences of real-life events) and their exposure formats (e.g., audio recordings, computer screens, videos, live staging) have varied widely, a common thread is that the items that measure true and false memories are neither novel nor unusual. The earlier birthday party example is a case in point. Birthday parties are customary occurrences in most children’s lives, and the targets and related distractors that were used to illustrate true and false memory (*chocolate* or *strawberry ice cream*, *blue* or *red birthday hat*, *lemonade* or *punch*) are likewise familiar to children. Analogously, in developmental studies in which the target materials were word lists, sentence lists, or narratives, the materials were chosen so that even young children would be well acquainted with all of the information that was presented (e.g., Ackerman, 1992, 1994; Paris & Carter, 1973; Reyna & Kiernan, 1994, 1995).

There is nothing odd about this. On the contrary, it is routine in all forms of developmental research to restrict attention to materials whose meanings are understood by children. In fact, any

study that violates this maxim is open to the automatic criticism (cf. Metzger et al., 2008) that the results may be uninterpretable, for a variety of reasons (e.g., that younger children will be confused, inattentive, or frustrated by unfamiliar information). Consequently, it is standard practice to use materials whose meanings are familiar to subjects of all ages (Ceci & Bruck, 1993). When it comes to false memory, however, FTT says that this habitual design feature has the unintended consequence of masking increases in false memory by reducing or eliminating age variability in the mechanism that generates such errors (e.g., Brainerd & Reyna, 2007). To take an example from a common type of study, suppose that the target material is a word list containing words such as *dog*, *yellow*, *toast*, *shirt*, and so forth, all of which are familiar exemplars of equally familiar taxonomic categories. Even young children are well aware that *dog* is an animal, *yellow* is a color, *toast* is a breakfast food, and *shirt* is an article of clothing (e.g., Bjorklund, 2004). In FTT, as we saw, it is gist memories of these concepts that support false recognition or false recall of related distractors such as *cat*, *green*, *cereal*, and *pants*. For these errors to increase with age, the tendency to store such memories would have to increase with age. By design, however, there is very little chance of this. Therefore, the major factor that will control age variability in false memory is increasingly good verbatim memory for exactly which exemplar of each category was on the list—*dog*, not *cat*; *yellow*, not *green*; *toast*, not *cereal*; *shirt*, not *pants*. This principle, that the influence of age variability in verbatim memory on age variability in false memory is maximized, while the influence of age variability in gist memory is minimized, is applicable to any experiment in which the target materials are familiar and meaningful to subjects of all ages.

### *Gist Sensitivity and Age Increases in False Memory*

Returning to the second question above, whether there are simple design modifications that would deliver tests of age increases in false memory, there is one that implements a well-known property of memory development that is concerned with the formation of semantic relations. As illustrated above, a hallmark of everyday domains of experience is that they contain multiple exemplars of familiar meanings: You will observe many different types of automobiles while waiting for a bus; you will type many different men's and women's names while answering your e-mail; you will try on many different types of coats, sweaters, and shirts while shopping for a winter wardrobe; and you will sit on many different types of sofas and recliners while shopping for furniture. A classic principle of memory development is that there is a lag of several years between the age at which children understand the meaning of a target (say, that Mary is a woman's name) and the age at which they readily connect that meaning across multiple exemplars (e.g., Alice, Barbara, Elizabeth, Judy, Mary, Rachel, Sally) to form semantic relations among those targets.

This principle plays a central role in predicting age increases in false memory, and to avoid any suggestion of circularity, it is important to establish that (a) meaning connection is defined and measured independently of false memory and (b) developmental trends in meaning connection have been verified independently of (and prior to) developmental reversals in false memory. On the former point, it is traditional to distinguish three levels of meaning connection for the verbal materials that dominate memory re-

search—namely, word–word, sentence, and story. The first level, which occurs with the most common type of research materials, word lists, consists of semantic relations between pairs of words. Recently, Wu and Barsalou (2007) developed a taxonomy that encompasses a myriad of word–word semantic relations but that consists of just six basic types of relations: antonymy, entity, introspective, situational, synonymy, and taxonomy. The individual relations that are grouped within each category are summarized in Table 1. When a memory experiment involves word lists (or pictures of corresponding words), lists can be readily scored for the presence of the pairwise relations in Table 1, and those scores predict aspects of memory performance (e.g., Cann, McRae, & Katz, 2006). With more complex materials, such as spoken or written discourse, the other two levels of meaning connection can be measured (e.g., Kintsch, Welsch, Schmalhofer, & Zimny, 1990). Sentence-level connection refers to overarching semantic relations among the words of a sentence (e.g., *The cage is on the table.*), whereas story-level connection refers to overarching relations that span a series of sentences (e.g., *The cage is on the table. The bird is in the cage. The cat is under the table.*). Following Bransford and Frank's (1971) influential work, children's ability to connect meaning at the sentence level is measured by administering unrepresented sentences that paraphrase studied sentences (e.g., *The table is above the cat.*), while their ability to connect meaning at the story level is measured by administering unrepresented sentences that combine information from different sentences (e.g., *The cage is above the cat.*). In either case, the children's task is to decide whether test sentences are consistent with the meaning of target materials (Reyna & Kiernan, 1994, 1995).

Table 1  
*Summary of Wu and Barsalou's (2007) Taxonomy of Semantic Relations Between Words*

Semantic relation	Definition
Antonymy	Words have opposite values of some meaning property (e.g., cold–hot, up–down).
Entity	One word names an entity (e.g., <i>chair</i> , <i>shirt</i> ) and the other names a property of that entity. Standard entity properties are the following: external components, external surface features, internal surface features, entity behaviors, quantities, systemic features, larger wholes, and spatial relations.
Introspective	One word names a mental state (e.g., <i>fright</i> , <i>happiness</i> ) and the other names a property that is related to that state. Standard introspective properties are the following: emotion, evaluation, contingency, representation, quantity, and negation.
Situational	Both words are related to the same situation (e.g., cooking, medical treatment). Standard situational relations are the following: function, action, participant, location, origin, time, manner, and associated entity.
Synonymy	Words have the same value of some meaning property (e.g., couch–sofa, trash–garbage).
Taxonomy	Words are related to the same taxonomic category (e.g., animals, colors). Standard taxonomic relations between words are the following: superordinate, subordinate, individual, and coordinate.

As mentioned, meaning connection is more limited in younger than in older children or adults, with the vast literature on age changes in discrimination shifts providing one of the earliest illustrations of this pattern (for a review, see Esposito, 1975). With respect to the three types of meaning connection that were just described, all of them improve with age. At the word–word level, there are dozens of developmental studies of one of the six categories of semantic relations in Table 1, taxonomic relations. The consistent finding has been that spontaneous formation of taxonomic relations between the targets on word or picture lists improves with age (Bjorklund, 1987, 2004; Bjorklund & Hock, 1982; Bjorklund & Jacobs, 1985; Bjorklund & Muir, 1988). Clustering in free recall of categorized lists has been by far the most popular method of measuring such improvements, and some typical data are exhibited in Figure 1b.

The plotted data are from an experiment by Bjorklund and Jacobs in which children of four age levels studied exemplars of four familiar categories (animals, occupations, seasonings, and weapons), with the exemplars being presented either randomly or blocked by category. Children's recall protocols were scored for the extent to which output was clustered by category, and as can be seen in Figure 1a, clustering improved with age. With respect to the other two levels of meaning connection, some typical age trends are displayed in Figure 1a. The plotted data are from an article by Reyna and Kiernan (1994) in which 6- and 9-year-olds studied three-sentence vignettes (e.g., the cat and bird vignette, above) and then responded to tests for recognition of sentence-level and story-level meaning. As can be seen, both forms of meaning connection roughly doubled between the ages of 6 and 9.

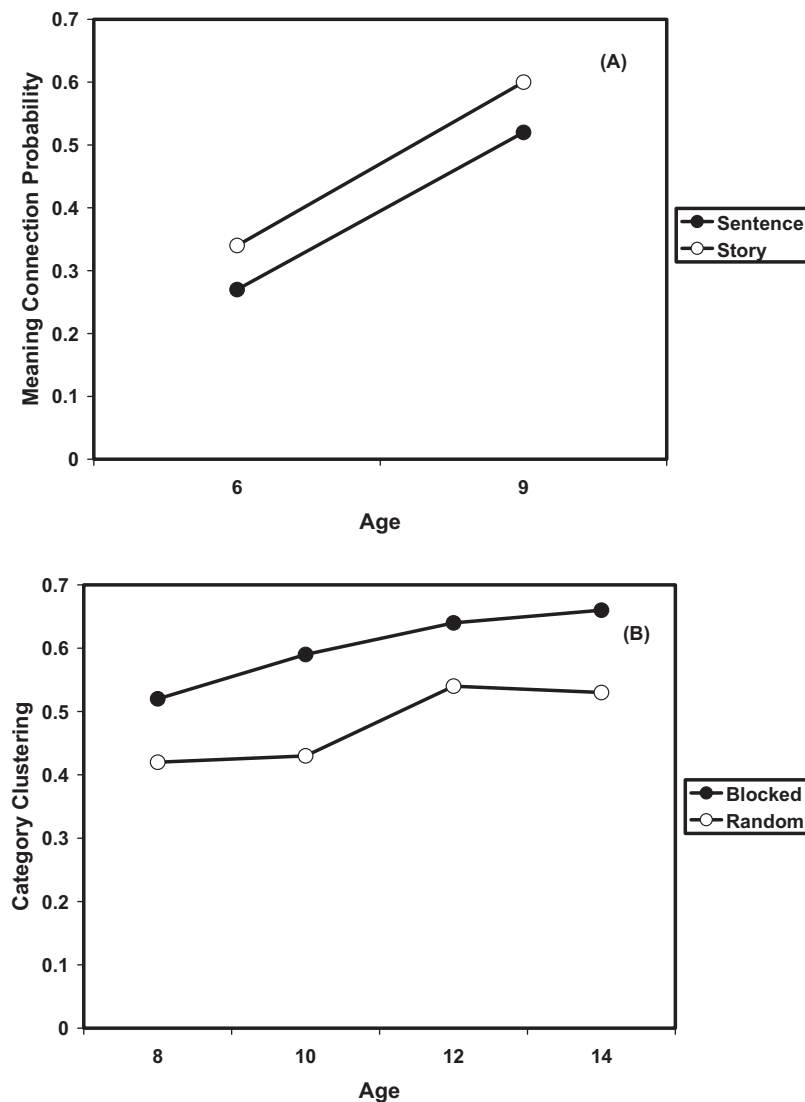


Figure 1. Developmental improvements in meaning connection (age in years). Panel A illustrates improvements at the sentence and story levels, using data from an article by Reyna and Kiernan (1994). Panel B illustrates improvements in connection at the word–word level, using data from an article by Bjorklund and Jacobs (1985).



When measures of meaning connection are administered to adults, rather than children, spontaneous formation of all three types of semantic relations is pervasive (e.g., Kintsch et al., 1990), which brings us back to false memory. In adults, the most powerful memory illusions are ones that emulate the everyday circumstance of encountering within a particular context multiple items that share meaning, with that context being the focus of later memory tests. Although there are many paradigms that have this feature, the most extensively studied ones are the Deese/Roediger/McDermott (DRM; Deese, 1959; Roediger & McDermott, 1995) illusion and false memory for categorized materials (e.g., Brainerd & Reyna, 1998b; Dewhurst, 2001; Tussing & Greene, 1999). In the DRM illusion, subjects study short lists of words (usually, 12–15), all of which are forward associates of a missing word (e.g., *window* or *sweet*). These lists are remarkably dense in word–word semantic relations between missing words and list words. Cann et al. (2006) and Brainerd, Yang, Howe, Reyna, and Mills (2008) scored the 55 DRM lists in the Roediger, Watson, McDermott, and Gallo (2001) norms for the presence of the six semantic relations in Wu and Barsalou's (2007) typology. Their results are shown in Table 2. The middle column displays the mean number of missing-word/list-word pairs per list that exhibit each of these relations, and the sum of the six values (14.62) is the average number of pairs per list that exhibits one of the six relations. Because there are 15 pairs per list (15 list words + 1 cue word), the value of 14.62 means that, on average, the members of every pair on every list are connected by one of these semantic relations. When adults study and recall such a list, the missing word intrudes in a large percentage of protocols (more than 50% for some missing words). If recognition tests are administered following several such lists, false-alarm rates for missing generating words are very high (70%–80% for some missing words) and, surprisingly, are usually as high as hit rates for targets.

In categorical false memory, on the other hand, subjects study a list of words or pictures, composed of blocks of targets that belong to the same taxonomic category—for example, 10 animals, 10 foods, 10 items of clothing, and 10 musical instruments (e.g., Howe, 2006; Seamon, Luo, Schlegel, Greene, & Goldenberg, 2000). The meaning connections in such lists are less dense than for DRM lists inasmuch as only some of the word pairs share meaning relations—namely, those that are exemplars of the same category—and only one of the

six semantic relations in Table 1 is involved. On recall tests, some categorically related but unrepresented exemplars of the categories are likely to intrude (usually, in 30%–50% of the protocols). On recognition tests, categorically related but unrepresented exemplars produce high false-alarm rates, sometimes approaching hit rates (e.g., Koutstaal et al., 2003; Koutstaal, Schacter, & Brenner, 2001).

Experimentation has pointed to two factors as being responsible for the high levels of false memory that are induced by connected-meaning tasks (for a review, see Brainerd & Reyna, 2005). First, when target materials allow subjects to form interitem semantic relations, this produces strong gist memories that are especially likely to be retrieved on memory tests, supporting false recall and recognition of related distractors (Reyna & Lloyd, 1997). Second, when target materials contain many exemplars of a given meaning, it becomes difficult to use verbatim traces of those exemplars to perform recollection rejection (Brainerd, Reyna, et al., 2003). In the categorized list example, for instance, being able to vividly remember *piano*, *trumpet*, *guitar*, *trombone*, and *clarinet* is not an indubitable basis for suppressing recall or recognition of an unrepresented musical instrument, such as *drums*, because subjects who connected the meaning across the musical instrument exemplars know that many exemplars were presented, the bulk of which are not coming to mind at the moment (Gallo, 2004).

This brings us back to developmental reversals in false memory. Brainerd et al. (2002) pointed out that connected-meaning tasks are promising test beds for the prediction that some types of false memory could increase dramatically with age, for two reasons. The first is children's well-documented limitation in forming semantic relations (e.g., Chi & Ceci, 1987), which entails that they will be less likely to store the strong gist memories that dominate adults' performance. The other reason is that age improvements in verbatim memory for targets will not provide older subjects with much of an advantage when it comes to suppressing intrusions and false alarms. Returning to the categorized list example, the fact that an adult can vividly remember *piano*, *trumpet*, *guitar*, *trombone*, and *clarinet* while a 7-year-old can only vividly remember *trumpet* and *clarinet* does not necessarily provide the adult with an advantage when it comes to suppressing *drums*, even though verbatim memory is two-and-a-half times better. Beyond forecasting developmental reversals, these distinctions generate numerous predictions about two groups of manipulations that affect the magnitude of the reversals that are ultimately observed (see Testing Process Models of Developmental Reversals). One group, *sufficiency manipulations*, consists of variables that make it easier for subjects to form semantic relations among targets and, hence, should increase false memory. Those increases should be less marked in older children (who are better at forming meaning connections on their own), which means that they ought to reduce age increases in false memory by elevating errors more in younger children than in older children. The other group, *necessity manipulations*, consists of variables that either interfere with subjects' ability to form semantic relations among targets or make it easier for them to use verbatim traces of targets to suppress false memories. Obviously, such manipulations should reduce false memories, but the reductions ought to be more marked in older children (who are more apt to form meaning connections and to preserve verbatim traces of targets). Consequently, necessity manipulations should reduce age increases in false memory by decreasing errors more in older than in younger children.

Table 2  
Numbers of Semantic Relations Between Target Words and Missing Words in Each of the 55 DRM Lists That Were Normed by Roediger, Watson, McDermott, and Gallo (2001)

Semantic relation	Statistic		
	Maximum number	Mean number	Standard deviation
Antonymy	3	0.44	0.76
Synonymy	7	1.02	1.53
Entity relations	13	2.65	3.03
Introspective relations	10	1.65	2.35
Situational relations	12	5.22	3.17
Taxonomic relations	12	3.64	3.15

### Review of Data: Three Lines of Evidence

We now examine developmental studies in which the prediction that connected-meaning tasks will produce age increases in false memory has been tested. This is the first of two sections in which the findings of such studies are reviewed. In the next section of this article, we consider findings for manipulations that were designed to test specific hypotheses about the underlying processes that control age trends in false memory. In the present section, we confine attention to results that bear narrowly on the age trends themselves. Results from three types of connected-meaning paradigms are reviewed, in separate subsections: (a) developmental studies of the DRM illusion, (b) developmental studies of false memory for categorized materials, and (c) developmental studies of memory suggestion.

#### *Developmental Reversals in the DRM Illusion*

In the DRM paradigm, it will be remembered, subjects are exposed to short word lists, all of which are forward associates of a missing word, followed by recall or recognition tests. The well-replicated adult patterns are high intrusion and false-alarm rates for these missing words (for a comprehensive review, see Gallo, 2006). This paradigm has generated the most extensive archive of studies demonstrating age increases in false memory. Below, such studies are reviewed, but first we briefly discuss a methodological issue, list length, that figures throughout the review.

#### *List Length*

Although we consider various methodological issues in connection with different studies, the length of the DRM lists that were administered is always noted because adult levels of false recall and false recognition vary directly with list length, while levels of true recall vary inversely with list length (e.g., Robinson & Roediger, 1997). In developmental studies, there is, of course, an understandable tendency to administer simpler tasks than have been used in adult research, but administering shorter DRM lists reduces sensitivity to age increases in the illusion. The power of a design to detect such increases shrinks as list length decreases because shorter lists constrict the range of variability in the DRM illusion at the high end (i.e., in adults; Brainerd et al., 2002). Opponent-processes distinctions provide a straightforward explanation of why this is so (see Reyna et al., 2007): Shorter lists supply subjects with fewer opportunities to make the meaning connections that foment errors and make it easier for subjects to retrieve the verbatim traces that suppress errors, which means that length is a necessity manipulation.

#### *Initial Uncertainty*

Developmental work on the DRM paradigm began with controversy when two articles appeared within 2 months of each other that seemed to produce contradictory findings. Brainerd et al. (2002) reported three experiments in which intrusion and false-alarm rates for missing words increased substantially with age, whereas Ghetti, Qin, and Goodman (2002) reported an experiment in which false-alarm rates for these same words did not vary with age, and intrusion rates were said to decrease with age. Brainerd et

al. (2002) reported two free-recall experiments and a recall-plus-recognition experiment. In the first recall experiment, 5-year-olds listened to a total of 10 DRM lists (12 words per list), and in the second, 5- and 7-year-olds listened to 16 lists (15 words per list). In contrast to adults' high intrusion rates, 5-year-olds falsely recalled the missing word only 5% of the time, and 7-year-olds recalled it only 7% of the time. The third experiment measured both false recall and false recognition of missing words. Five-year-olds, 11-year-olds, and young adults (undergraduates) listened to 16 DRM lists and responded to both free-recall tests and recognition tests (on which the 16 missing words appeared as related distractors, along with further probes for targets and unrelated distractors). The developmental patterns for false recall and false recognition are displayed in Figure 2, where the plotted data are the false recall probabilities of the missing words for the various age levels and the  $A'$  values for false alarms to missing words. The false recall probability more than tripled, from .11 to .37, across this age range. (The true recall probability increased as well, from .21 to .63.) The  $A'$  value for false recognition increased from .75 to .85 (while the corresponding value for true recognition increased by about the same amount). In sum, the overall developmental picture was (a) near-floor false recall in the youngest children, (b) significant false recognition in the youngest children, and (c) increases in both false recall and false recognition between early childhood and young adulthood.

A different picture was obtained by Ghetti et al. (2002), who reported that false recall decreased with age and false recognition was age invariant: "There were significant age differences in recall (5-year-olds evinced more false memories than did adults) but not in recognition of critical lures" (Ghetti et al., 2002, p. 705). However, Reyna et al. (2007) showed that the first half of this conclusion does not follow from Ghetti et al.'s data. In their study, 10 DRM lists of 7 words each (i.e., much shorter than the usual 12–15 words) were presented to 5-year-olds, 7-year-olds, and young adults, with the subjects responding to both free-recall and recognition tests. Analyses of variance showed that (a) true recall and true recognition both increased with age but (b) neither false recall nor false recognition varied with age. Finding b is inconsistent with the statement that 5-year-olds exhibited more false recall than adults did. Support for that statement was obtained from alternative scores that were generated by dividing the number of missing words that subjects recalled by the total number of words that were recalled. Reyna et al. pointed out that such scores selectively inflate false recall for subjects whose true recall is poorest, which happens to be the 5-year-olds:

This measure will yield higher false recall scores in younger children if true recall improves with age and false recall does not change because the same numerator (false recall) would be divided by a bigger and bigger denominator . . . Thus, it is not accurate to conclude from the alternative scores analysis that false recall is decreasing with age when it is actually not changing. (Reyna et al., 2007, p. 488)

In short, the difference between the Brainerd et al. (2002) and Ghetti et al. (2002) findings is that errors increased with age in the former case, while there were null age effects in the latter case. Reyna et al. (2007) pointed out that this difference is easily explained: There was little room for age increases in the Ghetti et al. study because the levels of false memory displayed by their adults were abnormally low, as compared with adult norms (e.g.,

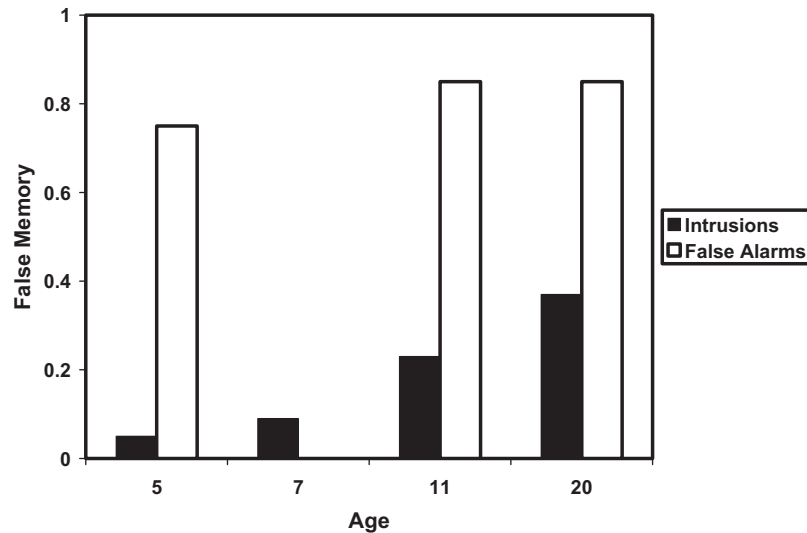


Figure 2. Intrusions of missing words at four age levels (in years) in the Deese/Roediger/McDermott illusion and false alarms to missing words at three age levels, based on experiments reported by Brainerd, Reyna, and Forrest (2002). The false-memory measures are intrusion probability for intrusions and  $A'$  for false alarms.

Stadler, Roediger, & McDermott, 1999). In Figure 2, Brainerd et al.'s adults falsely recalled missing words 37% of the time and falsely recognized them 85% of the time, but the corresponding levels in Ghetti et al.'s adults were approximately half of these values. The key question is why was the illusion so weak in Ghetti et al.'s adults? The likely answer, Reyna et al. concluded, is near-ceiling levels of true recall, which were probably caused by the administration of such short lists. DRM lists exhibit strong negative correlations between true recall, on the one hand, and false recall and false recognition, on the other (Gallo, 2006). Ghetti et al.'s adults recalled 89% of the target words, which is a far higher level than with DRM lists of standard length (Stadler et al., 1999). Thus, the developmental patterns in the Ghetti et al. (2002) and Brainerd et al. (2002) articles are not contradictory. Brainerd et al.'s findings confirmed the opponent-processes prediction that the DRM illusion should increase with age (because it is a connected-meaning illusion and age improvements in verbatim memory do not confer much of an advantage in editing out errors). Ghetti et al.'s study simply did not provide sensitive tests of this prediction (because their procedures reduced adults' error rates to such low levels that there was little room for age improvement).

### Subsequent Certainty

Any residual doubt about this conclusion has evaporated in the face of an avalanche of subsequent studies demonstrating age increases in the DRM illusion. The studies in question, which are summarized in Table 3, appear in articles by Anastasi and Rhodes, (2008); Brainerd, Holliday, and Reyna (2004); Brainerd, Forrest, Karibian, and Reyna (2006); Carneiro, Albuquerque, Fernandez, and Esteves, (2007); Dewhurst and Robinson (2004); Howe (2005, 2006, 2007, 2008); Howe, Cicchetti, Toth, and Cerrito (2004); Howe, Gagnon, and Thouas (in press); Lampinen, Leding, Reed, and Odegard (2006); Metzger et al. (2008); Sugrue and Hayne, (2006); and Weekes, Hamilton, Oakhill, and Holliday (2008).

These 15 articles contain a total of 22 experiments, all of which produced evidence of developmental reversals in the DRM illusion. When these articles are combined with the Brainerd et al. (2002) and Ghetti et al. (2002) articles, the published archive consists of 26 experiments, 25 of which detected age increases in false recall or false recognition of missing words. As can be seen in Table 3, the age range encompassed by these experiments is preschool through the early 20s, and the lengths of DRM lists varied from 7 to 15 words. Beyond their consistent verification of the developmental reversal prediction, the subsequent articles add six findings whose combined effect is to broaden the scope of the developmental reversal effect considerably. We describe each of these findings separately.

*Developmental reductions in net accuracy.* We speculated that, counterintuitively, the net accuracy of recall or recognition might sometimes be found to decline with age. When a paradigm produces age increases in false memory, it is possible, in principle, for net accuracy (the ratio of true memory to false memory plus true memory) to be lower in adults and older children than in younger children. Some of the articles in Table 3 report this pattern. Brainerd, Forrest, Karibian, and Reyna (2006) found that the net accuracy of recall declined between the ages of 6 and 14; Howe (2008) found that the net accuracy of recall declined between the ages of 5 and 11; Lampinen et al. (2006) found that the net accuracy of recognition declined between age 8 and young adulthood; and Metzger et al. (2008) found that the net accuracy of both recall and recognition declined between age 7 and young adulthood.

*List length.* We mentioned that shorter lists reduce adults' false memories, while increasing their true memories, and that, therefore, shorter lists reduce power to detect age increases in the DRM illusion. As can be seen in Table 3, some experiments used longer lists and others used shorter lists, but one (Sugrue & Hayne, 2006) directly compared age trends in false memories for longer

Table 3  
*Published Developmental Studies of the Deese/Roediger/McDermott Illusion*

Study	Age span (years)	List length	Memory test	Result
Anastasi et al. (in press) <sup>a</sup>	5–adult	15	Recall + recognition	Age increase in false recognition and recall
Brainerd et al. (2002)	5–adult	12	Recall + recognition	Age increase in false recognition and recall
Brainerd et al. (2004)	7–14	11	Recognition	Age increase
Brainerd et al. (2006) <sup>b</sup>	6–14	12	Recall + recognition	Age increase in false recognition and recall
Carneiro et al. (2007) <sup>a</sup>	3–adult	8–15	Recall + recognition	Age increase in false recognition and recall
Dewhurst & Robinson (2004)	5–11	8	Recall	Age increase
Ghetti et al. (2002)	5–adult	7	Recall + recognition	No age change
Howe (2005)	5–11	14	Recall	Age increase
Howe (2006)	5–11	14	Recall	Age increase
Howe (2007)	8–12	12	Recall + recognition	Age increase in false recall
Howe (2008)	5–11	12	Recall	Age increase
Howe et al. (2004)	6–11	12	Recall + recognition	Age increase in false recognition and recall
Howe et al. (in press)	6–adult	14	Recall + recognition	Age increase in false recognition and recall
Lampinen et al. (2006)	6–adult	7	Recognition	Age increase
Metzger et al. (in press) <sup>a</sup>	7–adult	8	Recall + recognition	Age increase in false recognition and recall
Sugrue & Hayne (2006)	5–adult	7 & 14	Recall + recognition	Age increase with 14 words but not with 7
Weekes et al. (2008) <sup>c</sup>	9–11	10	Recall + recognition	Ability difference in false recognition and recall

<sup>a</sup> The Carneiro et al. (2007), Metzger et al. (2008), and Anastasi and Rhodes (in press) experiments all used child-normed DRM lists. Metzger et al. and Anastasi and Rhodes also used adult-normed lists. False recognition and false recall increased with age with both types of lists.

<sup>b</sup> The second experiment in Brainerd et al. (2006) included samples of children with learning disabilities, who displayed smaller age increases in false memory than children without learning disabilities.

<sup>c</sup> Weekes et al. (2008) compared children who had been classified with semantic-processing disability to children who had not been so classified and found that false memory was lower in children with disabilities.

versus shorter lists. That experiment was a recognition-plus-recall design that paralleled Brainerd et al.'s (2002) Experiment 3 in the sense that subjects of different ages (5- and 21-year-olds) studied and recalled a series of DRM lists, followed by a recognition test composed of target probes, related distractors (missing words and other semantically-related words), and unrelated distractors. The lists consisted of 14 words for half the subjects and 7 words for the other half. The results confirmed the differential sensitivity prediction about shorter versus longer lists. The intrusion probability increased significantly with age (from .37 to .50, interpolating from Sugrue and Hayne's Figure 1) for longer lists but not shorter ones, and the false-alarm probability for missing words also increased significantly with age (from .58 to .80, interpolating from Sugrue and Hayne's Figure 2) for longer lists but not shorter ones. As of the present writing, this is the only published experiment providing direct evidence that shortening DRM lists masks age trends in false memory by selectively suppressing error rates at older age levels.

*Special populations.* In forensic applications of research on false memory, an abiding question is whether findings are comparable for children from populations that are represented disproportionately in legal cases, such as maltreated children, children from low-income homes, and children who are learning disabled or who have retardation (e.g., Bruck & Ceci, 1999). Developmental reversal findings for the DRM illusion have been extended to such populations by Howe et al. (2004), Brainerd et al. (2006), and Weekes et al. (2008). Howe et al. (2004) examined the strength of the illusion in samples of maltreated children, children from low-income families, and control children. They found that intrusions and false alarms to missing words increased between the ages of 8 and 11 in all three groups of children and by equivalent amounts. Brainerd et al. (2006) and Weekes et al. explored developmental

reversals with two definitions of "development," the usual chronological age definition and another involving learning ability. Brainerd et al. (2006) reported a study in which half of the children at each of two age levels (7 and 11 years) had been classified as learning disabled and the other half had not. The prediction, based on opponent-processes models, was that both definitions would reveal developmental reversals; that the DRM illusion would not only be more marked for older than for younger children but would be more marked for nondisabled than for disabled children. (The second half of this prediction was derived from prior research, e.g., Swanson, 1991, showing that learning-disabled children's ability to form semantic relations is more limited than nondisabled children's.) The prediction was confirmed. Among nondisabled children, false recall of missing words increased from .16 in 7-year-olds to .30 in 11-year-olds, and within both age levels, false recall was lower for disabled than for nondisabled children. Similarly, Weekes' et al. found that the DRM illusion is reduced in learning-disabled children. They studied a more specific disability that ought to have direct consequences for connecting meaning across DRM list words: reading comprehension deficits that are due to impairments in semantic processing. Weekes et al. conducted a recall-plus-recognition experiment in which lists were administered to children who were classified as exhibiting a semantically-based reading comprehension deficit and children who were not classified as reading disabled. As predicted, both false recall and false recognition of missing words were lower in the disabled children.

*Developmental reversals with child-normed lists.* The research that has been mentioned thus far relied on lists that were generated from adult word-association norms (Roediger et al., 2001; Stadler et al., 1999). Metzger et al. (2008), Anastasi and Rhodes (in press), and Carneiro et al. (2007) speculated that



developmental reversals might disappear if DRM lists were constructed from child word-association norms because children may fail to understand some of the words that are derived from adult norms, rather than fail to connect meanings that they understand. To evaluate this possibility, Metzger et al. generated association norms for 8-year-olds and young adults; Anastasi and Rhodes generated norms for 5- and 8-year-olds; and Carneiro et al. generated norms for 4-year-olds, 7-year-olds, 12-year-olds, and young adults. Metzger et al. administered both child- and adult-normed DRM lists to 7-year-olds, 10-year-olds, and young adults; Anastasi and Rhodes administered both child- and adult-normed lists to 5-year-olds, 8-year-olds, and young adults; and Carneiro et al. administered lists to 4-year-olds, 7-year-olds, 12-year-olds, and young adults that had been normed for their respective age levels. The key finding in all instances was that the developmental reversal pattern was present with child-normed as well as adult-normed lists. For instance, Metzger et al. found that, between age 7 and young adulthood, false recall rose from .02 to .16 and false recognition from .23 to .72 with adult-normed lists but that false recall rose from .01 to .13 and false recognition rose from .28 to .46 with child-normed lists. Similarly, Anastasi and Rhodes found that between age 5 and young adulthood, the probability of false recall increased from .23 to .33 with adult-normed lists and from .12 to .33 with child-normed lists as well as that the probability of false recognition increased from .45 to .67 with adult lists and from .36 to .53 with child lists. Thus, contrary to the hypothesis that age increases in false memory are manufactured by adult-normed lists, Anastasi and Rhodes's young children displayed lower levels of error for the same missing words when the lists were child-normed. Finally, by using DRM lists that had been separately normed for each age level, Carneiro et al. also found that false recall and false recognition increased with age. In sum, the DRM illusion has been shown to increase with age for child-normed as well as adult-normed lists.

*Developmental reversals in other languages.* In most of the articles in Table 3, DRM lists were presented to subjects whose native language was English. However, developmental reversals have also been detected in other languages. For instance, Carneiro et al.'s (2007) experiments were conducted in Portuguese, and Howe et al. (2008) reported a study with 6-, 8-, and 12-year-old children and young adults who were bilingual in English and French. DRM lists were presented in French for half the subjects at each age level and in English for the other half of the subjects. After each list, subjects performed a free-recall test in either French or English, and after the last list, subjects responded to a recognition test composed of target probes, related distractors (the missing word for each list and other unrepresented words that were semantically related to each list), and unrelated distractors. Regardless of whether lists were studied and recalled in English or in French, false recall of missing words increased across the four age levels and so did false recognition of missing words. However, both types of errors were more common in English than in French. Some interesting findings were obtained by comparing unilingual testing conditions (list presentation and memory tests in the same language) to bilingual testing conditions (list presentation in one language and memory tests in the other language). False memories were higher in 5-year-olds when modalities matched than when they mismatched; match–mismatch did not affect 8- and 12-year-olds' levels of false memory; and false memories were lower in

young adults when modalities matched than when they mismatched. The match–mismatch effect in adults is consistent with adult findings for other modality manipulations (for a review, see Reyna & Lloyd, 1997).

*Picture lists.* In most of the articles in Table 3, DRM lists were presented as spoken or written word lists. However, Howe (2008) has conducted an extensive program of research in which DRM lists were presented as line drawings or color photographs. This research followed up an earlier finding (see below) that picture lists produce lower levels of false memory for categorized materials than word lists produce. In an initial experiment, the subjects were 5-, 7-, and 11-year-olds, who studied and recalled lists that were presented as words or as line drawings or as color photographs. Intrusions of missing words increased with age, from 12% in the youngest children to 31% in the oldest. There was also a surprising finding about presentation modality. Although, as expected, line drawings suppressed false recall at all age levels, relative to words, color photographs did not. Additional experiments were undertaken to determine why, contrary to expectation, color photographs did not suppress false recall. We will return to these experiments in the next section, when we review results that bear on processes that underlie developmental reversals in false memory. For now, the point is merely that the developmental reversal effect has been confirmed for picture lists as well as the usual word lists.

### *Summary of Developmental Patterns in the DRM Illusion*

We have reviewed 17 articles that contain 26 experiments in which false recall or false recognition of the missing words of DRM lists was measured in subjects that ranged in age from 3 years to early 20s. The experiments provide massive evidence of an age-increase pattern for intrusion and false-alarm rates. Owing to design differences among these experiments, such as the variable of list length, it is not possible to be specific about the exact amounts of increase between particular age levels. However, if we stipulate that a standard DRM list is one consisting of 12–15 words, the modal result has been for false recall to roughly triple between age 5 and early adolescence and for false recognition to roughly double.

These patterns confirm theoretical predictions and disconfirm the notion that age decreases in false memory are universal trends. Two other patterns are instructive. First, as mentioned at the outset, the standard view of memory development is that net memory accuracy increases steadily between early childhood and young adulthood. While that must be true if false memory always decreases with age, it becomes an empirical question if false memory can also increase with age. As we saw earlier, FTT allows for the counterintuitive possibility that the net accuracy of memory may sometimes decline with age (because, with particular tasks, age increases in intrusions or false alarms may outstrip age increases in correct recall or hits). Although declines in net accuracy were not inevitable in the experiments that we reviewed, there was clear evidence of them in some experiments. Thus, it appears that this surprising result, which is anticipated on theoretical grounds, sometimes occurs.

The other instructive pattern is that robust developmental growth in false memory has been detected for both intrusions and false alarms. The generality of age increases across recall and

recognition is theoretically probative because it suggests that the trend is storage-driven—that developmental changes in the information that is extracted from DRM targets are implicated. This rests on the fact that from the standpoint of retrieval, recall and recognition are very different types of tests. According to FTT, recognition tests provide better retrieval cues for verbatim traces than free-recall tests do, and hence, other things being equal, recognition is more slanted toward reading out the information that is stored in verbatim traces than free recall is because recognition tests re-present targets (cuing verbatim traces of those targets; e.g., Seamon et al., 2002). Because the retrieval environments on the two types of tests are so different, variables whose effects are retrieval-driven rather than storage-driven typically affect free recall and recognition in different ways. This is not true for age changes in false memory for missing words, which suggests that these changes are storage-driven (see Kimball, Smith, & Kahana, 2007, for a mathematical model that shows that this assumption is necessary to produce good fits to developmental DRM data). That robust age increases in false memory are obtained for both recall and recognition is also of forensic importance. Questions that involve yes–no recognition and questions that involve recall are both used in investigative interviews of children (e.g., Poole & Lamb, 1998), and it has often been reported that levels of false memory are lower for recall (e.g., Sternberg et al., 2001), which is a retrieval effect, of course. Consequently, it might be thought that age increases in false memory could be eliminated or greatly reduced by relying on recall tests. Developmental DRM data obviously do not support such a conjecture.

### *Developmental Reversals in False Memory for Categorized Materials*

We turn now to studies that have identified age increases in false memory by using a second type of connected-meaning procedure—namely, the presentation of categorized materials. It will be remembered that, in this paradigm, the responses that count as false memories are intrusions or false alarms involving items that are either unrepresented exemplars (e.g., *pants*, *blouse*) or unrepresented labels (e.g., *clothing*) of presented exemplars. The target materials consist of multiple exemplars of one or more categories (e.g., *belt*, *coat*, *dress*, *scarf*, *shirt*, *socks*, *shoes*, *skirt*, *tie*) followed by recall tests, on which missing exemplars and category labels may crop up as intrusions, or by recognition tests, on which the related distractors are missing exemplars and/or category labels. Below, we review developmental studies that have demonstrated age increases in such intrusions and false alarms, using two different procedures.

#### *Categorized Picture Recognition*

Sloutsky and Fisher (2004b) used a powerful memory-falsification procedure that they had developed in earlier experiments (Sloutsky & Fisher, 2004a). In this paradigm, the target materials exemplify the same familiar categories (e.g., domestic cats), but targets are so similar to each other that, as in the DRM illusion, it is difficult to use verbatim traces of them to suppress false alarms to unrepresented exemplars. FTT predicts that such procedures should produce high levels of false memory in adults (they do; see Koutstaal et al., 2003; Koutstaal & Schacter, 1997)

and, therefore, that they should also produce robust developmental increases in false memory. Sloutsky and Fisher's (2004b) subjects were 5-year-olds and 19-year-olds. The target materials consisted of color photographs of 10 cats, 10 bears, and 10 birds. There were three presentation conditions, one of which was designed to foment false memories. In that condition, as pictures were presented, the subjects were provided with a verbal cue that differentiated the cats ("has beta cells inside its body") from the bears and the birds ("does not have beta cells inside its body"). Picture presentation was followed by a recognition test consisting of target probes (presented pictures of cats and bears), related distractors (unpresented pictures of cats), and unrelated distractors (unpresented pictures of squirrels). True and false memory both increased during this 14-year age span. The false-alarm rate for related distractors nearly doubled, from .41 to .76, while the hit rate rose from .72 to .83. It appears from these numbers that false memory increased more than true memory. To test this difference for statistical significance, Sloutsky and Fisher (2004b) computed  $A'$  values that indexed subjects' ability to discriminate targets from related distractors. They found that, indeed, the mean  $A'$  value was significantly larger in children (.66, interpolating from Sloutsky and Fisher's, 2004b, Figure 2) than in adults (.52).

Fisher and Sloutsky (2005) reported more extensive developmental studies of the procedure that Sloutsky and Fisher (2004b) first used. One of them used the same procedures as those just described, except that the subject sample allowed for more extensive age comparisons. This time, the procedure was administered to samples of 5-, 7-, 11-, and 19-year-olds (undergraduates). The developmental pattern for false memory was as follows: There was a slight increase in false alarms between the ages of 5 and 7 (from .40 to .45) that was not reliable; there was a large increase between the ages of 7 and 11 (from .45 to .59) that was reliable; and there was another large increase between the ages of 11 and 19 (from .59 to .74) that was reliable. The developmental picture for true memory was somewhat different: The hit rate increased reliably between the ages of 5 and 7 (from .70 to .77), between the ages of 5 and 11 (from .70 to .79), and between the ages of 5 and 19 (from .70 to .81), but none of the other age comparisons was reliable. Thus, once again, net accuracy declined between early childhood and young adulthood. As before, to evaluate this possibility Fisher and Sloutsky computed  $A'$  values that indexed subjects' ability to discriminate targets from related distractors. They found that, indeed, memory accuracy again declined with age. The mean  $A'$  values for 5- and 7-year-olds (.71 and .75, interpolating from Fisher and Sloutsky's Figure 2) did not differ significantly. However, there was a significant drop between the ages of 7 and 11 (from .75 to .63), and there was another significant drop between the ages of 11 and 19 (from .63 to .54). Fisher and Sloutsky reported three further experiments with this procedure that did not focus on developmental increases in false memory, and hence, we do not consider them here.

Most recently, further developmental data have been reported on Sloutsky and Fisher's (2004b) procedure by Wilburn and Feeney (in press), who conducted two experiments. In Experiment 1, they replicated the condition in which Sloutsky and Fisher (2004b) first identified age increases in false recognition of unrepresented pictures that exemplified presented categories. The subjects were 5-year-olds and young adults. Like Sloutsky and Fisher (2004b), Wilburn and Feeney found that the false-alarm rate for related

distractors increased with age but by a smaller amount (from .46 to .51). However, unlike Sloutsky and Fisher (2004b), the hit rate for targets decreased substantially between early childhood and young adulthood (from .72 to .62). Consequently, like Sloutsky and Fisher (2004b), the  $A'$  measure of subjects' ability to discriminate presented and unpresented exemplars decreased markedly during this age range (from .68 to .54). A second experiment with subjects from the same two age levels produced similar results: The false-alarm rate for related distractors increased with age (from .49 to .57); the hit rate decreased with age (from .72 to .65); and hence, the  $A'$  measure of subjects' ability to discriminate presented and unpresented exemplars decreased with age (from .68 to .56). Wilburn and Feeney's data agree with Sloutsky and Fisher's (2004b) on two key points—namely, that between early childhood and young adulthood, false memory for unpresented exemplars of presented categories increases and net accuracy declines. The only notable difference between the two sets of experiments is that true memory increased with age in Sloutsky and Fisher's (2004b) research but decreased with age in Wilburn and Feeney's. To isolate possible reasons for this difference, we compared the true- and false-memory performance of 5-year-olds and young adults across the two sets of experiments. That comparison showed that Wilburn and Feeney's 5-year-olds performed comparably with Sloutsky and Fisher's (2004b) on both the true- and false-memory measures. However, Wilburn and Feeney's young adults performed more poorly than Sloutsky and Fisher's (2004b) on both types of measures. Thus, when it comes to recognition memory, the ability levels of the two adult samples were not the same.

### *Categorized Word Lists*

Brainerd et al. (2004), Howe (2006), and Chaing and Chiu (2006) all reported experiments in which children of different ages studied and recalled or recognized words lists of exemplars of familiar taxonomic categories. Except for the target materials, the Brainerd et al. (2004) experiment followed the same procedure as developmental DRM experiments. Specifically, children of different age levels (5- and 11-year-olds) listened to three lists, followed by a recognition test, followed by two more cycles of listening to three lists and responding to a recognition test. Each list consisted of 10 familiar exemplars of 10 familiar categories (e.g., animals, clothing, colors). On the recognition test, probes for targets, related distractors (missing exemplars and category labels), and distractors were presented. The mean  $d'$  value for missing exemplars and category labels increased across this age range, from 0.80 to 1.15.

Howe (2006) obtained larger age increases in false memory in three experiments in which 5-, 7-, and 11-year-old children studied and recalled categorized lists of words or lists of pictures that depicted the objects named by the corresponding words. In Experiment 1, each child listened to eight 14-item lists on which all the items belonged to a single taxonomic category (e.g., animals, body parts) and performed a free-recall test immediately following the presentation of each list. False recall of unpresented exemplars of these categories nearly doubled during this age range (from .17 to .28 to .32, interpolating from Howe's, 2006, Figure 1). This procedure was repeated in a second experiment, and false recall roughly doubled during this age range (from .17 to .30 to .38). In a third experiment, developmental increases in false memory interacted with presentation modality. As mentioned earlier in con-

nection with the Ghetti et al. (2002) study, developmental increases are not possible if the levels of false memory in older subjects are forced to abnormally low levels. Presenting DRM lists as pictures, rather than as spoken or printed words, is known to have this effect on adults' false recall of missing words (Hege & Dodson, 2004). When Howe (2006) repeated the procedure of the first two experiments, substituting categorized lists of black-and-white line drawings for categorized lists of words, he obtained a similar finding. Specifically, when 5-, 7-, and 11-year-olds studied and recalled these pictorial lists, false recall of unpresented exemplars was suppressed well below the levels that had been observed for words in 11-year-olds (.20 vs. .38, interpolating from Howe's, 2006, Figure 3) and 7-year-olds (.19 vs. .30, interpolating from Howe's, 2006, Figure 3), but in 5-year-olds, the levels of false recall were the same for pictures and words.

Finally, Chaing and Chiu (2006) reported an experiment in which 9- and 11-year-olds listened to four 12-word lists on which all the words belonged to a single taxonomic category (e.g., animals, musical instruments). The design resembled Howe's (2006) in that the children performed a free-recall test after they had listened to each list. After they had recalled the fourth list, however, a recognition test was administered on which the probes were targets, related distractors (unpresented exemplars from the presented categories), and unrelated distractors. As compared with the research that has been reviewed thus far, the inclusion of such a narrow age range, coupled with the relatively small numbers of subjects that participated at each age level ( $n = 24$ ) and the small number of lists per subject (four), provided less power to detect age changes. Consistent with these design limitations, the key overall finding was no age trends: Neither false recall nor false recognition of missing exemplars varied significantly between age levels, and neither true recall nor true recognition varied significantly between age levels.

### *Summary of Developmental Trends in False Memory for Categorized Materials*

Developmental studies of categorized materials and developmental studies of the DRM illusion have yielded convergent findings on the prediction that under connected-meaning conditions false memories will increase between early childhood and young adulthood. Except for a study by Chaing and Chiu (2006) that raises power questions, false memory for categorized materials has provided consistent evidence of age increases in false memory for unpresented items that preserve the meaning of familiar taxonomic categories. This pattern has been detected in multiple studies using two different procedures: Sloutsky and Fisher's (2004a) categorized pictures methodology and categorized word lists.

Another similarity between developmental DRM studies and developmental studies of categorized materials concerns age trends in net accuracy. In DRM studies, net accuracy sometimes decreased with age. This is a theoretically important result because FTT anticipates that the net accuracy could decline with age under specified conditions. Although there was some evidence of such age declines in DRM studies, one of the two paradigms discussed in the present section has consistently produced declines in net accuracy. Beginning with Sloutsky and Fisher's (2004b) initial study, every experiment that implemented this procedure has found that adults are less able than young children to discriminate

unpresented exemplars from presented exemplars. In this one paradigm, the counterintuitive age trend in net accuracy that is anticipated by FTT has been the norm.

Another instructive similarity between DRM studies and studies of categorized material is that developmental growth in false memory has been obtained for both intrusions and false alarms. Sloutsky and Fisher's (2004a) paradigm provides strong evidence of such growth for false recognition of missing exemplars, while the categorized-word-list procedure provides correspondingly strong evidence for false recall of missing exemplars. As we remarked earlier, that age increases in false memory hold for both recognition and recall is of considerable theoretical importance because it suggests that these trends are storage-driven, rather than merely being retrieval effects. Further, as we also remarked earlier, that age increases in false memory hold for both recognition and recall is of considerable forensic significance because recall and recognition questions both figure in investigative interviews of children, and it is widely thought that free recall is especially unlikely to provoke false reports. The developmental data on categorized materials, like the developmental DRM data, do not support that contention.

### *Developmental Reversals in Memory Suggestion*

Counterintuitive results can be contagious. Once they are discovered in one domain, they are apt to show up in other places where they would not otherwise have been anticipated. That is the current situation in developmental research on memory suggestion. Since Ceci, Ross, and Togliola (1987) reported a series of experiments showing that susceptibility to memory suggestion decreases between early childhood and young adulthood, numerous studies have been published in which Loftus's (1975) classic misinformation procedure was used to implant false memories in children of various ages. Following an initial period of uncertainty over whether age declines in susceptibility to suggestion were reliable (for a review, see Ceci & Bruck, 1993), Ceci et al.'s (1987) age trend has been replicated again and again. The robustness of this finding has led to its routine use in expert testimony, where it is regularly opined that children's memories are inherently more falsifiable than those of adolescents and adults (e.g., Ceci & Bruck, 1995). Such expert testimony has become so commonplace that judges sometimes rule against admitting it on the ground that this point is already well understood by jurors and, therefore, expert testimony is not needed to educate them on it (McAuliff, Nicholson, & Ravanshenas, 2007).

Once sufficient data had been reported to establish the age-increase pattern for the DRM illusion and categorized materials, however, developmental studies began to appear that detected the same pattern for memory suggestion. This literature, being very recent, is far less extensive than developmental work on the other two paradigms. However, some studies have now confirmed the pattern, demonstrating that under some conditions, older children are more vulnerable to memory suggestion than younger children. These reverse developmental trends have sometimes been connected to age differences in children's knowledge or "gist processing" (Brainerd & Reyna, 1995), such that higher levels of knowledge are associated with heightened susceptibility to suggestion. For example, in an experiment that preceded the recent intensification of research on age increases in suggestibility, Ceci, Caves,

and Howe (1981) reported that the usual age-decrease pattern did not hold: Seven- and 10-year-olds were equally susceptible to suggestions about cartoon characters for whom their knowledge levels were equivalent.

In what follows, we focus on four studies in which age increases in susceptibility to memory suggestion were identified by deliberately manipulating the same variable as in the developmental studies of the DRM illusion and categorized materials—namely, meaning connectedness of the target events. Then, we summarize some experiments by Pezdek and Roe (1995, 1997) that produced early evidence of age increases in susceptibility to suggestion.

### *Connected-Meaning Studies*

Connolly and Price (2006) reported an experiment that implemented the connected-meaning variable in a play session in which their child subjects participated. The authors proposed that it might be possible to produce age increases in suggestibility by administering memory suggestions about groups of events with strong meaning connections. They implemented this idea by having preschool and 1st-grade children participate in either a single play session or four play sessions. The multisession condition was of principal interest because meaning connectedness was manipulated in that condition. During each play session, children participated in eight activities. During each activity, they experienced a critical detail (e.g., a white object). When that activity was repeated during the next play session, children experienced a new critical detail that preserved key meaning features of the previous detail (e.g., a brown, blue, or yellow object). For some of the eight activities, the level of meaning association among the various details was higher than for the details that were used for other activities. A misinformation session followed the last play session, during which the memory suggestions focused on critical details that the children had not experienced but that preserved the meanings of the details they had experienced (e.g., a black object, when the colors white, brown, blue, and yellow had been experienced). One day later, the children responded to free- and cued-recall tests about the events they had experienced during the play sessions. The predicted reversal in suggestibility was obtained in the connected-meaning condition. For memory suggestions about critical details whose meanings were very strongly associated, the intrusion rate for suggested details tripled across this age range on free-recall tests, while it quadrupled on cued-recall tests. For memory suggestions about critical details that were less strongly associated, the intrusion rate for suggested details more than doubled across this age range on free-recall tests, while it tripled on cued-recall tests. In short, although age decreases in suggestibility have been a dominant outcome in this literature for many years, a single experiment that deliberately implemented conditions that are specified in FTT's model of false memory produced sizeable age increases in susceptibility to suggestion.

An experiment by Ross et al. (2006) used the same logic to produce age increases in susceptibility to suggestion in an especially important type of memory error in criminal cases: false eyewitness identifications of innocent suspects. Most research on child memory suggestion is concerned with distortions in the accuracy of children's descriptions of events that they have witnessed or participated in. However, children may be called on to provide eyewitness identifications of suspects. It is well estab-



lished that adults' identifications are highly susceptible to suggestions that occur after crimes have been witnessed (for a review, see Brainerd & Reyna, 2005). Although there is little experimentation that is directly concerned with the suggestibility of children's eyewitness identifications, most expert witnesses opine that children's identifications are more susceptible to suggestion than adults' (Kassin, 2001). However, Ross et al. reasoned that (a) older subjects would be more likely than younger ones to process the meaning relations between culprits and innocent bystanders (e.g., age, gender, ethnicity) and, therefore, (b) to falsely identify bystanders as culprits. To evaluate this hypothesis, children of four age levels (5-, 7-, 9, and 11-year-olds) viewed a video of interactions between preschool teachers and children. In the control condition, one event depicted a female teacher reading a story to the children, while in the suggestion condition, the same event depicted a male teacher reading the story. At the end of the video, in both conditions, a female teacher was shown entering a cafeteria, sitting down at a table beside a man, placing her wallet on the table, and then leaving to go to a vending machine. While her back was turned, the man removed money from her wallet and left the cafeteria. Later, the children were administered an eyewitness identification test that was composed of five photographs, the innocent bystander (the male teacher who read the story in the suggestion condition) and four foils. Children were told to identify the thief's picture, and they were told that the thief's picture might or might not be in the photo spread. For children in the suggestion condition, the probability of falsely identifying the innocent bystander more than tripled over this age range, from .18 (5-year-olds) to .40 (7-year-olds) to .47 (9-year-olds) to .64 (11-year-olds). Also, the net accuracy of identification performance, defined as the correct response rate divided by the false identification rate plus the correct response rate, declined markedly with range. Thus, like Connolly and Price's experiment, a theoretically-motivated misinformation design produced age increases in suggestibility.

The next article, by Ceci et al. (2007), implemented another variation on the connected-meaning theme to produce age increases in suggestibility. Two experiments were reported: a judgment experiment that established variations in the semantic distance among 22 common objects for 4- and 9-year-old children and a suggestibility experiment with children of the same two age levels in which the semantic distance between targets and related distractors was manipulated. In the first experiment, children were presented with pictures of common objects (e.g., a lemon, an orange, a glass of milk, a cow). The pictures were presented in triads, and for each triad, children were asked to select a picture that did not belong with the other two. The resulting data were subjected to multidimensional scaling to generate a semantic space for these objects, for both 4- and 9-year-olds. The analysis yielded a series of semantic distance scores for each object, which measured the strength of the meaning relation between that object (e.g., cow) and the other objects (e.g., milk, horse, lemon). Ceci et al. (2007) pointed out that according to FTT, subjects should be more susceptible to suggestion when the semantic relation between targets and related distractors (memory suggestions) is strong than when it is not. Experiment 2 was a memory suggestion study in which children were first exposed to a picture story in which 8 of the rated objects were presented as part of the story, followed by an interview that included suggestions about 4 of these objects, followed by a recognition test containing probes for targets, sug-

gested objects, and unrelated distractors. The semantic distance analysis was used to determine whether children's susceptibility to suggestion depended on the strength of the semantic relation between targets and suggested objects. It did, as expected on theoretical grounds. However, crucially, they also predicted age increases in suggestibility as a function of semantic relatedness. To test that prediction, they identified pairs of targets and suggested items for which the semantic relation between those objects was stronger in 9-year-olds than in 4-year-olds (e.g., lemon/orange, egg/cheese). The prediction, of course, was that susceptibility to suggestion should increase with age for such pairs. The data showed that false alarms to the related distractors in such pairs roughly doubled across this age range. The standard age-decrease effect was obtained for other pairs of targets and related distractors. Thus, opposite age trends in susceptibility to suggestion were produced under theoretically specified conditions.

The last study, by Fazio and Marsh (in press), reports an experiment that detected increases in suggestibility across three age levels: 5-, 6-, and 7-year-olds. In an earlier experiment, Marsh, Meade, and Roediger (2003) discovered that adults are susceptible to false suggestions about real-world facts when those facts are embedded in target materials with very strong meaning relations and the facts fit well with those relations, such as false facts embedded in short stories that revolve around familiar themes (e.g., a vacation cruise, a couple's 50th wedding anniversary, the Civil War). On cued-recall tests, young adults recalled false facts one-quarter of the time after reading such stories once and recalled false facts 40% of the time after reading such stories twice. Fazio and Marsh hypothesized, based on FTT, that susceptibility to such suggestions would increase with age. On its face, this is a counterintuitive prediction: Adults and older children are more likely to know true facts than young children and, hence, should be better able to resist suggestions. However, Fazio and Marsh reasoned that if susceptibility to suggestion depends on forming strong gist memories of the meaning relations in stories, susceptibility ought to increase with age. Fazio and Marsh created four stories that were more child appropriate than Marsh et al.'s: a trip to the Eiffel Tower, a trip to the Leaning Tower of Pisa, some farmyard animals' search for the sun, and a skunk who learned to defend himself. Each child listened to two of the stories and then responded to a cued-recall test for all four stories. The key outcome was that the developmental trend in false memories followed theoretical predictions rather than intuitive expectations based on age differences in fact knowledge. The tendency to recall false facts (e.g., that autumn is another word for spring, that there are 10 eggs in a dozen) was higher in 7-year-olds than in younger children: The mean probabilities of recalling false suggested facts were .06 (5-year-olds), .05 (6-year-olds), and .13 (7-year-olds).

### *The Pezdek and Roe Experiments*

The results that were just reviewed are all theory-driven in the sense that the investigators used the memory principles that we discussed earlier to formulate designs that could detect age increases in susceptibility to memory suggestion. However, more than a decade ago, two developmental misinformation experiments were reported by Pezdek and Roe (1995, 1997) that were not motivated by these theoretical ideas but that happened to include the connected-meaning features that are deemed to be important in

producing age increases in susceptibility to suggestion. Those experiments, too, revealed that older children were more susceptible to memory suggestions than younger children.

In their first experiment, Pezdek and Roe's (1995) subjects were 4- and 10-year-old children. The children viewed two picture stories, one about a man working at a construction site and one about a woman in a kitchen preparing to bake a cake. With respect to connected meaning, the first story presented multiple exemplars of familiar construction items, such as tools (e.g., hammer, saw), while the second presented multiple exemplars of familiar kitchen items, such as baking utensils (e.g., spoon, bowl). After viewing the picture stories, the children received a postevent interview that contained misinformation about some of these items (e.g., that the man had used a screwdriver, rather than a hammer; that the woman had used a fork, rather than a spoon). This was followed by a recognition test composed of target probes, related distractors (misinformation), and unrelated distractors. Large age increases in susceptibility to suggestion were detected. Over this age range, the mean  $d'$  value for false memory (the tendency to accept false information that had been suggested over false information that had not been suggested) increased from  $-0.03$  to  $2.13$  for items that had appeared once in the stories (e.g., the hammer, the spoon), and it increased from  $0.08$  to  $2.53$  for items that had appeared twice in the stories.

The developmental patterns were analogous in Pezdek and Roe's (1997) second experiment. As before, the age range of the subject sample was 4 to 10 years. The target materials and suggestions were different, however. The children participated in a 25-minute activity, part of which consisted of viewing a picture story. During the picture story, the experimenter touched the child's hand for 10 s, or touched the child's shoulder for 10 s, or did not touch the child. Fifteen minutes later, children were given a postevent interview that contained misinformation about the touching. Some of the children who were touched on the hand were told that they had been touched on the shoulder, and some of the children who were touched on the shoulder were told that they had been touched on the hand. At the end of the session, the children responded to a recognition test that included target probes and probes for misinformation about touches. The mean  $d'$  value for false memory for a suggested touch when children had been touched in a different location doubled with age, from  $0.72$  in 4-year-olds to  $1.49$  in 10-year-olds.

### *Summary of Developmental Reversals in Memory Suggestion*

Although research on age increases in suggestibility is more recent and therefore thinner on the ground than the other two lines of investigation, all three converge on two key conclusions. First, under conditions that, theoretically, ought to make memory increasingly susceptible to distortion with age, spontaneous distortions and distortions that originate in misinformation have both been found to intensify with age. That this trend has been confirmed for implanted as well as spontaneous false memories is critical because, as others have pointed out (e.g., Toglia, Ross, Ceci, & Hembrooke, 1992), a plethora of social persuasion factors (e.g., coercion, source prestige, trust in others) are operative in misinformation research that are absent in the other two lines of investigation. This fact, as Ceci and Bruck (1993) discussed in

their review of child-suggestibility research with children, has prompted some writers to propose that misinformation-induced errors are not false memories but, rather, are merely social demand effects. Although the subsequent literature has demonstrated that such errors are indeed false memories (Brainerd & Reyna, 2005), persuasion factors contribute to the power of misinformation to foment false memories. The additional fact that younger children are more susceptible to persuasion than older children or adults (Ceci et al., 1981; Toglia et al., 1992) has led some to propose that this is responsible for the traditional finding of greater susceptibility to memory suggestion among younger children (Ceci & Bruck, 1993). If so, it would be impossible to reverse this trend under theoretically favorable conditions (i.e., connected meaning), because younger children's heightened sensitivity to social persuasion would offset the effects of age increases in the ability to process meaning relations. However, the studies that we reviewed show that it is possible to reverse the traditional age trend, and therefore, the indicated conclusions are that notwithstanding younger children's heightened sensitivity to persuasion, the traditional trend is not immutable and implanted false memories will increase with age under theoretically-specified conditions.

The second point of convergence among developmental studies of suggestibility, the DRM illusion, and categorized materials concerns the generality of the age-increase trend over recognition and recall tests. As was the case for the other two lines of research, age increases in susceptibility to suggestion have been detected for both intrusions and false alarms. The theoretical significance of this fact is the same as before. That is, developmental changes on the storage side of memory would seem to be implicated because recall and recognition tests provide children with very different retrieval cues, with recognition tests providing better retrieval cues for verbatim traces of targets than recall tests provide. The fact that, despite the radically different retrieval environments on recall and recognition tests, age increases in suggestibility have been identified with both types of tests suggests that something more than developmental changes in retrieval are responsible.

### Testing Process Models of Developmental Reversals

It might be thought that the many demonstrations of age increases in false memory provide firm support for the theoretical ideas that forecast the increases in the first place. Actually, however, mere confirmation of age increases delivers no direct support for those ideas. To briefly recapitulate, the core notion is FTT's principle that children's limitations in forming meaning relations among targets imply that age increases in false memory ought to be detected in connected-meaning paradigms, especially if it is difficult to use verbatim traces of targets to suppress intrusions and false alarms. The logical relation between this principle and the age-increase pattern is one of implication: If the principle is valid, the pattern will be observed. If the pattern is observed, however, it does not follow that the principle is valid because other possibilities exist. The actual causes could be different than the posited mechanism.

How can the validity of this mechanism be tested, then? Validity turns on results that are produced by experimental manipulations of two types, *sufficiency* manipulations and *necessity* manipulations. Sufficiency manipulations answer a question about younger children, who show lower levels of false memory in connected-

meaning paradigms: Can age increases be shrunk by enhancing their ability to establish meaning relations among target events? Necessity manipulations, on the other hand, answer a question about older children, who show higher levels of false memory in connected-meaning paradigms: Can age increases be shrunk by interfering with their ability to establish meaning relations among target events or by making it easier for them to use their superior verbatim memories to suppress errors? In extremis, especially powerful sufficiency manipulations could increase younger subjects' levels of false memory to older subjects' levels, and especially powerful necessity manipulations could decrease older subjects' levels of false memory to younger subjects' levels. If data answer both the sufficiency and necessity questions affirmatively, a one-to-one relation has been established between age increases in false memory and FTT's explanation of such increases. If only the sufficiency question is answered affirmatively, which is a common occurrence in developmental research (e.g., Ceci et al., 2007; Glaser & Resnick, 1972), then there are additional mechanisms that produce the observed age increases. If neither the sufficiency nor the necessity question is answered affirmatively, mechanisms other than those posted in FTT are responsible for developmental reversals in false memory.

In the first two subsections that follow, we review findings for sufficiency and necessity manipulations. As an advance organizer, available data show that (a) manipulations that help children connect meaning across targets increase false memory in younger children, narrowing the age difference between younger and older children and sometimes eliminating it, and that (b) manipulations that eliminate or interfere with the formation of meaning relations decrease false memory in older children, again narrowing the age difference and sometimes eliminating it. In the third subsection, we summarize other findings that bear on the validity of FTT's explanation of developmental reversals but that do not involve either sufficiency or necessity manipulations. In the fourth subsection, we sketch some recent proposals by Howe (2005, 2006) that attempt to explain developmental reversals for two connected-meaning paradigms (the DRM illusion and categorized word lists) with somewhat different theoretical assumptions.

### *Results for Sufficiency Manipulations*

If developmental improvements in spontaneous meaning connection are behind the age increases that were reviewed earlier, a sufficiency manipulation is merely one that helps children to make meaning connections among targets that they otherwise might not make. Such manipulations ought to be more helpful to children who have difficulty forming such connections spontaneously than to children who do not. There is one manipulation of this ilk, gist cuing, that has been extensively investigated. Gist cuing was inspired by classic research on semantic organization in children's recall (e.g., Bjorklund, 1987) and consists of providing advance warnings about the presence of meaning relations and encouraging children to extract that information. There are two other sufficiency manipulations that have been studied in one experiment apiece: embedding targets in a narrative that makes their meaning connections obvious (Dewhurst, Pursglove, & Lewis, 2007) and administering targets for which young children are better at forming meaning relations (Ceci et al., 2007). We review findings on gist cuing first, followed by results for the other manipulations. As

will be seen, the results of all of these manipulations support the conclusion that the meaning connection mechanism is sufficient to produce age increases in false memory.

### *Gist Cuing*

At one time, gist-cuing was a central feature of research on children's recall of categorized lists, the objective being to induce children to display the organizational effects that are so prominent in adults' recall. A key datum, for other semantic-processing manipulations as well as for gist cuing, was an inverted-U relation between age and the effectiveness of these manipulations. Specifically, the manipulations tended to be more effective at some intermediate age level—usually, later childhood or early adolescence—than during early childhood or young adulthood. The explanation that emerged from experimentation (see Bjorklund, Miller, Coyle, & Slawinski, 1997) was that semantic-processing manipulations are maximally effective with children who are old enough to have already developed some basic semantic abilities, and they are less effective when children are so young that they lack most of these abilities or when children are so old that the development of these abilities is virtually complete. Extrapolating this well-established pattern to developmental studies of false memory, the expectation would be that gist cuing ought to be more effective at increasing false memories in somewhat older children than in very young children or adults. However, the exact age at which gist cuing is optimally effective depends on how powerful the manipulation is—how strong a prosthetic for meaning connection it provides—with more powerful manipulations being optimally effective at younger ages. For instance, the gist-cuing manipulations that are discussed in this subsection appear to be optimally effective during the mid-to-late elementary grades. However, a more powerful manipulation introduced by Dewhurst et al. (2007), which is discussed in the next subsection, appears to be optimally effective in kindergarten children.

The initial gist-cuing experiment was reported by Brainerd et al. (2004), and more recent ones have been reported by Brainerd et al. (2006); Howe (2006); Holliday, Reyna, and Brainerd (in press); and Lampinen et al. (2006). Brainerd et al.'s (2004) and Lampinen et al.'s (2006) experiments dealt with false recognition, whereas the others dealt with false recall. Taking recognition first, Brainerd et al. (2004) administered gist cues for categorized lists, and Lampinen et al. (2006) administered them for DRM lists. The basic methodology of Brainerd et al.'s (2004) experiment was described earlier. In addition, a condition was included in which children were told that the upcoming list words would all be exemplars of a specific taxonomic category (e.g., "all the words on the next list will be animals"). This manipulation elevated both younger and older children's false recognition of unrepresented exemplars, but the elevation was greater in younger children. Consistent with the notion that meaning connection is sufficient to produce age increases in false memory, there was no age increase when the gist of each list was stated before presentation. Lampinen et al.'s (2006) results for DRM lists were analogous. It will be remembered that these authors reported an experiment with 6-, 8-, and 20-year-olds in which false recognition of missing words increased substantially with age. They also reported an experiment that implemented the same basic procedure but that included a cuing condition to stimulate the processing of meaning relations.

In that condition, children were told (a) that all of the words on each DRM list go together and (b) that they were to think about how the words on each list go together as they listened to the list. The children were also given warm-up lists (e.g., the names of holidays) for which the between-target meaning relations were stated by the experimenter as the lists were presented. An analysis of  $d'$  values revealed that false recognition once again increased with age in a no-cues control condition. The more interesting results were for the gist-cuing condition. In line with the inverted-U relation that has been obtained in developmental studies of categorized recall, the cuing manipulation increased false memory more at the intermediate age level (8-year-olds) than at the youngest or the oldest age level. Consistent with the sufficiency hypothesis about meaning connection, age increases in false memory were smaller in the gist-cuing condition, and the increase between age 8 and age 20 was not reliable in the cuing condition.

Holliday et al.'s (2008) results were similar to Lampinen et al.'s (2006) inasmuch as they also detected an inverted-U relation between gist cuing and DRM false memory, though for intrusions rather than false alarms. Five age levels were studied: 7-, 9-, 11-, 13-, and 15-year-olds. All subjects listened to six 14-word DRM lists and responded to a separate free-recall test for each list. There were two conditions, a control condition and a gist-cuing condition. In the cuing condition, children were told, just before they listened to each list, that all the words would have very similar meanings, and a theme was stated for the upcoming list (e.g., "they are all medical words"). In the control condition, as usual, false recall of missing words increased dramatically with age, more than tripling between the ages of 7 and 15. Consistent with the sufficiency hypothesis about meaning connection, although false recall also increased with age in the gist-cuing condition, the increase was much smaller. Similar to Lampinen et al.'s (2006) results, an inverted-U relation was observed such that gist cuing increased the intrusions of 9- and 11-year-olds more than the intrusions of either 7-year-olds or 13- and 15-year-olds. In fact, gist cuing did not produce reliable increases in intrusions at the latter age levels.

Brainerd et al. (2006) also studied the effects of gist cuing on DRM false recall, and their data echoed those of Holliday et al. (2008) and Lampinen et al. (2006). Subjects at three age levels (6-, 11-, and 14-year-olds) listened to sixteen 12-word DRM lists and responded to a free-recall test for each list. As in Holliday et al. (2008), there was a control condition and a gist-cuing condition. The latter was somewhat different than Holliday et al.'s (2008) because the children participated in a pretraining phase before any of the DRM lists were presented. This consisted of three problems on which (a) the experimenter read 4 words aloud that were well-known exemplars of a familiar taxonomic category (e.g., animals) and (b) the children responded to a four-alternative multiple choice item that stated possible semantic relations between the items (e.g., animals, clothing, colors, furniture). The experimenter provided the correct answer following each choice. After the pretraining phase, the DRM lists were presented and recalled. The children were told that their task was to figure out how the words on each list were similar to each other and were reminded about this before each list was presented. In the control condition, intrusions of missing words increased with age in the usual way, from 6% to 26% between the ages of 6 and 11 and from 26% to 51% between the ages of 11 and 14. Consistent with the sufficiency hypothesis, age increases in false recall were reduced in the

gist-cuing condition. Although false recall was lower for 6-year-olds than for 11- or 14-year-olds in that condition, there was not a reliable increase between the ages of 11 and 14 because gist cuing was more effective in 11-year-olds than in 6- or 14-year-olds: Cuing increased the probability of missing word intrusions by .22 in 11-year-olds but by only .09 in 6-year-olds and .04 in 14-year-olds.

In the above studies, gist cuing produced results that support the sufficiency hypothesis about meaning connection. That is, age increases in false memory were attenuated by gist cuing, and age increases were sometimes eliminated. In contrast, one article has appeared that obtained null findings for gist cuing. Howe (2006) reported some research in which gist cuing was used with free recall of categorized lists by children of three age levels (5-, 7-, and 11-year-olds). Each child listened to a total of eight 14-word categorized lists, presented either as words or as pictures. Half the children at each age level participated in a control condition, and half participated in a gist-cuing condition. In the cuing condition, children were given the name of the taxonomic category to which all of the words on each list belonged before the list was presented. This form of cuing failed to produce reliable differences between the control and cuing conditions in either true or false recall. In the only other study of gist cuing with categorized lists, Brainerd et al. (2004) found that the same type of cuing elevated children's true and false memory for category exemplars. A possible explanation of these different findings lies in differential sensitivity of the memory tests: Brainerd et al. (2004) administered recognition tests, while Howe (2006) administered recall tests.

### Other Sufficiency Manipulations

Presenting targets in isolation places greater onus on rememberers to make meaning connections than when the targets are presented in sentences and narratives that spell out those connections. For instance, that the semantic relation "food" holds among the words *apple*, *bread*, *cheese*, *hamburger*, *milk*, and *salad* can be made clear by presenting these words as part of a story about children eating lunch. According to the sufficiency hypothesis, the general effect of such embedding ought to be the same as gist cuing—false memories should be elevated and age increases in false memories should be attenuated or eliminated. This prediction was tested by Dewhurst et al. (2007), using DRM lists. In a control condition, children of three ages (5-, 8-, and 11-year-olds) listened to 14-word DRM lists (the *cold*, *doctor*, *fruit*, *lion*, *music*, *sleep*, *smell*, and *thief* lists) and, following each list, responded to a recognition test. In a narrative condition, the same lists were read, except that the 14 words were presented in a series of sentences that told a story. For instance, the words for the *doctor* list were presented as follows:

The *nurse* had written a prescription for Sally because she was *sick*. Her mum, who was a *lawyer*, told Sally she had to take the *medicine* because it would improve her *health*. She said if Sally did not take it she would have to go to the *hospital*. Sally hated them more than the *dentist*. Sally saw a *physician* the last time she was *ill*. She went into his *office* and he listened to her heart with a *stethoscope*. She then went to a different *clinic* where she saw a *surgeon* who gave her the treatment she needed to *cure* her.

The results of Dewhurst et al.'s (2007) experiment resemble those of some of the gist-cuing experiments. On the one hand, false



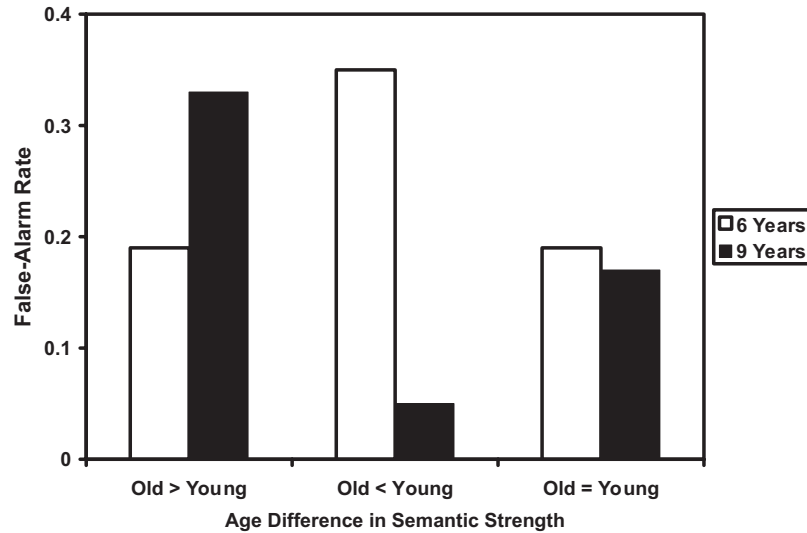


Figure 3. Age differences (in years) in suggestibility as a function of age differences in the strengths of the semantic relations between objects in a misinformation experiment reported by Ceci, Papierno, and Kulkořksy (2007).

recognition of missing words increased with age in the control condition. On the other hand, false recognition of missing words did not increase with age in the narrative condition. The absence of age increases in this condition was due to the fact that although narrative presentation greatly elevated the false-alarm rate in 5-year-olds (from .5 to .8), it did not affect the false-alarm rate in either 8- or 11-year-olds.

Turning to Ceci et al.'s (2007) study demonstrating age increases in susceptibility to memory suggestions, it will be recalled that 9-year-olds were more suggestible than 6-year-olds when the suggestions referred to objects whose meaning relations were stronger in older children. However, this can also be made into a sufficiency manipulation as follows. Under FTT's explanation of age increases in false memory, increases in suggestibility should disappear when suggestions refer to objects for which meaning relations are equally strong in younger and older children, and further, younger children should be more suggestible than older children when suggestions refer to objects for which meaning relations are stronger in younger children. In other words, the sufficiency hypothesis says that in this design, differences in subjects' susceptibility to suggestion should depend on which semantic relations are strongest for them. It can be seen in Figure 3 that this is precisely the pattern that was observed. The three pairs of bars in this figure show the false-alarm rates for suggested information when the underlying semantic relation was stronger in older children (left pair), was equally strong in younger and older children (right pair), and was stronger in younger children (middle pair). As predicted, the direction of the age change was determined by differences in the strengths of semantic relations.<sup>3</sup>

Actually, Ceci had obtained analogous findings some years before in an experiment that did not involve misinformation. In that experiment (Ceci & Howe, 1978, Experiment 2), 7- and 13-year-olds studied and recalled lists of familiar words, with the false-memory measure consisting of intrusions of unrepresented words that were consistent with the meanings of list words. How-

ever, the list words (e.g., *ball*) had two meanings, a dominant one (a play object, as in bat and ball) and a secondary one (a formal dance). The assumption was that younger children would be less likely than older children to encode words' secondary meanings but just as likely to encode their primary meanings. If so, the prediction, based on FTT, is an Age  $\times$  Meaning crossover—specifically, that older children should display higher intrusion rates than younger children for secondary meanings, but the reverse should be true for primary meanings. That is the pattern that Ceci and Howe obtained.

### Conclusion

Summing up, the results that have just been reviewed provide support for the hypothesis that age differences in meaning connection are sufficient to produce age increases in false memory. When manipulations made it easier for young children to connect meaning across targets, such age increases in false memory were reduced and sometimes they were completely eliminated. Manipulations ranged from cuing the gist of the upcoming word lists to providing narrative connections among related words.

### Results for Necessity Manipulations

The logic of necessity manipulations is the opposite of the logic of sufficiency manipulations: If age increases in false memory are

<sup>3</sup> Semantic distance and chronological age were both significant predictors of children's suggestibility. When generalized estimating equations (GEE) models were run, age and distance were significant when both variables were included in the model, which suggests that chronological age still accounted for variance after semantic distance was controlled. ROC curves indicated that although semantic distance had a greater area under the ROC curve than did age, the difference was not statistically reliable.

rooted in better spontaneous meaning connection, those increases ought to disappear if that advantage is completely excised, and they ought to be attenuated if that advantage is reduced (e.g., by interfering with opportunities to form meaning connections). We refer to the former as *elimination manipulations* and to the latter as *reduction manipulations*. Examples of both have been studied. Pertinent findings are summarized below, in separate subsections.

### *Elimination Manipulations*

Three methods of completely removing the contribution of older subjects' superior meaning connection to age trends in false memory have been investigated: (a) measuring false memory for related distractors when children are exposed to only one semantically-related target, rather than several; (b) measuring false memory for related distractors when children are exposed to several targets whose interconnection relation is not semantic; and (c) interfering with subjects' ability to form meaning connections between semantically related targets. Experiments implementing the first approach have been reported by Connolly and Price (2006) and Brainerd and Reyna (2007). Experiments implementing the second approach have been reported by Holliday and Weekes (2006) and again by Brainerd and Reyna (2007). An experiment implementing the third approach has been reported by Lampinen et al. (2006).

Concerning the first approach, as previously discussed, age increases in false memory were obtained by Connolly and Price (2006) and Brainerd and Reyna (2007) by measuring false memory for items whose meanings had been exemplified by multiple targets. Beyond this, each experiment included a control condition in which false memory for the same related distractors was measured but there was no opportunity for false memories to be based on connected meaning because only one semantically-related target had been presented (e.g., if the related distractor was *cat* or *animal*, the only related target was *dog*). In this no-connection condition, age increases in false memory disappeared. For instance, it will be remembered that in Connolly and Price's experiment, false recall of suggested details increased markedly with age when children had encountered four targets that were strongly related in meaning to those details. However, when children had encountered only one target that was strongly related in meaning, false recall of those same suggested details did not vary with age. Brainerd and Reyna's (2007) results, which are displayed in Panel A of Figure 4, were qualitatively similar. In their experiment, false recognition of category labels and unrepresented category exemplars roughly tripled, between the ages of 6 and 14, when children had been exposed to lists containing blocks of eight exemplars of those categories. However, on those same lists, some categories were represented by only a single target. The outcome, as can be seen in Figure 4, was opposite age trends. False recognition of related distractors increased with age when connected meaning could be used to support such errors, but false recognition of the same items decreased with age when connected meanings could not be formed.

Turning to the second elimination methodology, this technique, which involves switching from semantic relations to phonological ones, was developed by Holliday and Weekes (2006). If older subjects' advantages in forming meaning connections are responsible for age increases in false memory, increases should not be observed for false memories that are supported by nonsemantic

relations that even young children readily connect across targets. For example, when subjects respond to a recognition test on which related distractors sound like targets (e.g., the word *battle*, if the word *cattle* was a target), false-memory effects are observed for such items (e.g., Rotello, 2001), with the effects increasing in size as the number of sound-alike targets increases (Brainerd & Reyna, 2005). Because even young children readily notice when things sound alike, Holliday and Weekes reasoned that age increases would disappear if false memories were measured that are based on such phonological relations and, further, such false memories would decline with age. To test this conjecture, they conducted a phonological analogue of their DRM experiment, which we describe earlier. That is, rather than listen to 10-word DRM lists, 8-, 11-, and 13-year-old children listened to 10-word lists on which the targets sounded alike (e.g., *pole*, *hole*, *bowl*, . . . ) and responded to recognition tests on which the related distractors preserved these phonological relations (e.g., *coal*, *roll*). As predicted, the age trends in these phonological false memories were the reverse of the age trends that had been obtained for DRM lists. The false-alarm rate for related distractors declined from .42 (age 8) to .30 (age 11) to .29 (age 13), with the difference between 8- and 11-year-olds and 8- and 13-year-olds being reliable but not the difference between 11- and 13-year-olds.

Brainerd and Reyna (2007) replicated Holliday and Weekes' (2006) results, using related distractors that either sounded like eight targets that children had listened to or like only one target that they had listened to. The procedure paralleled the aforementioned method of varying connected meaning: For some related distractors, the list contained a block of eight sound-alike targets, and for other related distractors, the list contained only one sound-alike target. The results of this experiment can be seen in Panel B of Figure 4. As would be expected, the false-alarm rate for related distractors was higher when related distractors sounded like eight prior targets rather than one prior target, and this was true for 6-, 10-, and 14-year-olds. Crucially, however, the age trend in false memory for connected-phonology relations was the opposite of the usual age trend for connected-meaning relations. Thus, consistent with the necessity hypothesis about connected meaning, age increases in false memory disappeared when false memories were supported by surface relations among targets.

The final elimination methodology is a gist-interference technique that was introduced by Payne, Elie, Blackwell, & Neuschatz (1996). In the DRM illusion, the formation of semantic relations between the words on individual lists is encouraged by presenting all of the words on each list together, in a block. Payne et al. reasoned that the storage of gist memories of such relations by adults would be greatly impaired if subjects studied a single long list, composed of many DRM lists whose words were presented in random order. They reported findings that were consistent with this hypothesis. If it is true that young children are far less likely to form semantic relations between the words on a DRM list, even when they are presented together in a block, randomization should not interfere with this process as much in young children as it does in adults. On that basis, Lampinen et al. (2006) predicted that the usual age increase in the DRM illusion would be eliminated if the words on DRM lists were presented in random order rather than in blocked fashion. In their experiment, children of two age levels, 6- and 8-year-olds, and a sample of young adults were presented with a single list that

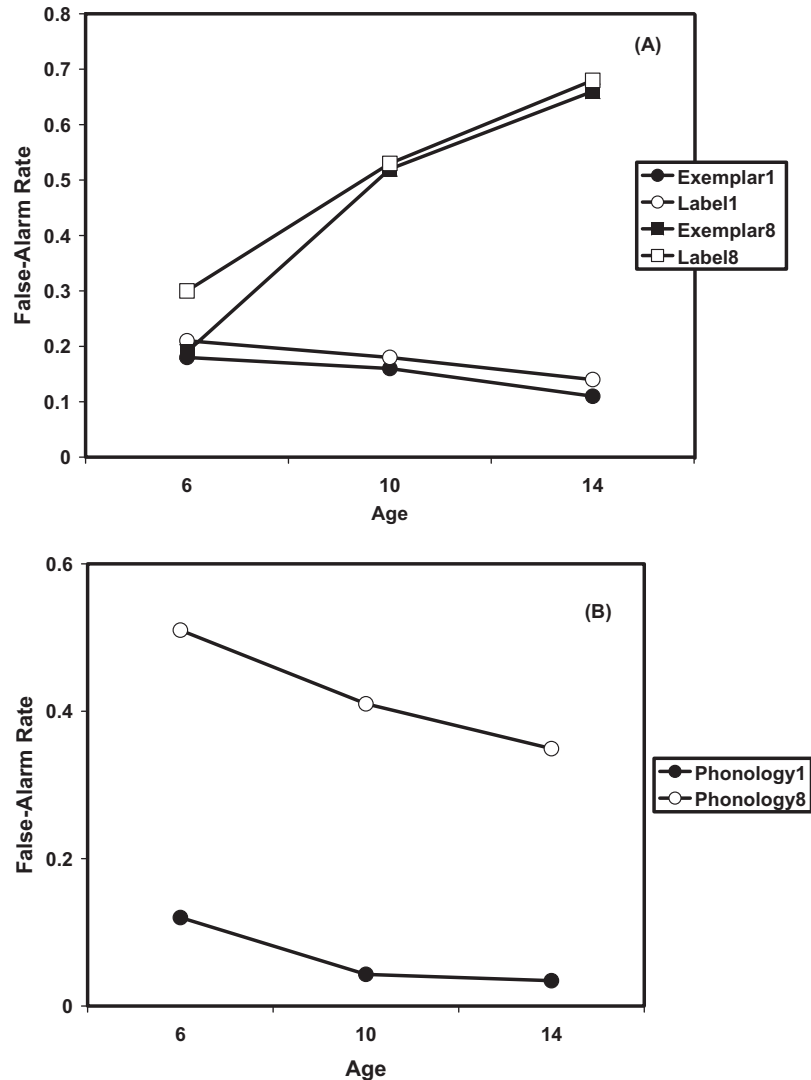


Figure 4. Opposite age trends (in years) in false memory for semantic versus phonological relations in an experiment by Brainerd and Reyna (2007). Panel A illustrates age trends for semantic false memory. Panel B illustrates age trends for phonological false memory.

was composed of seven DRM lists, followed by recognition tests. Presentation was blocked for half the subjects and random for half the subjects. The key data are shown in Figure 5, where  $A'$  values for missing words are plotted. Note that consistent with prediction, the usual age increase in false recognition (blocked condition) vanished under random presentation.

#### Reduction Manipulations

The logic of these manipulations is that although false memories are measured that can be based on connected meaning, targets are presented under conditions that make meaning connection more difficult, thereby reducing or even eliminating the advantage that older subjects normally have in this area. It follows, under the necessity hypothesis about the role of meaning connection in false memory, that such manipulations should attenuate age increases in false memory or even eliminate them. Currently, data are available

on three such manipulations: list length (Sugrue & Hayne, 2006), list format (Howe, 2006; 2008), and list content (Carneiro et al., 2007; Metzger et al., 2008).

If older children and adults are better at meaning connection, their advantage over younger children increases as the number of semantically-related targets increases (Brainerd & Reyna, 2005). That is, semantic relations that support false memory will become stronger as more exemplars of those relations are encountered, if subjects readily connect meaning across targets but not if they fail to do so. Thus, according to the necessity hypothesis about meaning connection, the difference between the levels of false memory that are observed for younger versus older children should increase as the number of prior exemplifications of meanings increases. Connolly and Price's (2006) and Brainerd and Reyna's (2007) techniques of exemplifying a meaning once versus multiple times is the limiting case of this manipulation. A less extreme version

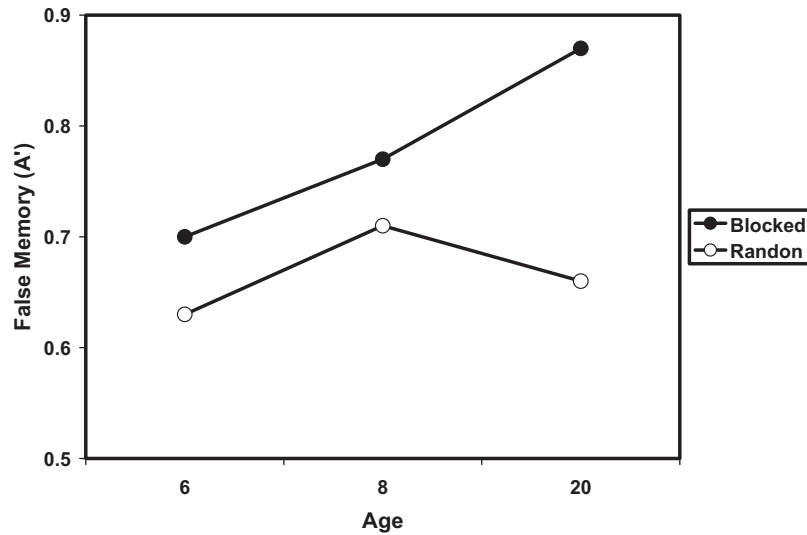


Figure 5. Demonstration by Lampinen, Leding, Reed, and Odegard (2006) that the Deese/Roediger/McDermott (DRM) illusion increases with age (in years) when DRM lists are presented in blocked order but not when they are presented in random order.

consists of comparing age trends in false memory when a meaning has been exemplified by a small number of targets versus a large number. That, it will be recalled, was the nature of Sugrue and Hayne's (2006) experiment. Specifically, 5-year-olds and 20-year-olds studied and recalled DRM lists and also responded to recognition tests. Half the lists consisted of 14 words and half consisted of 7 words. The longer lists produced the usual age increases in intrusion and false-alarm rates for missing words, but the shorter lists did not. Consistent with the necessity hypothesis about meaning connection, the reason that age increases disappeared with shorter lists was the same for intrusions and false alarms—more specifically, that shorter lists suppressed 20-year-olds' memory errors much more than they suppressed young children's.

The next necessity manipulation takes advantage of the fact that in adult memory research, certain target presentation formats are known to direct processing away from semantic content and, hence, to hinder the formation of semantic relations among targets. For instance, if the targets are a list of words, presenting the list visually (printed words) leads to less semantic processing than presenting it orally (Smith & Hunt, 1998), and if the list is presented visually, printing the words in several unusual fonts leads to less semantic processing than printing them in one familiar font (Arndt & Gould, 2006; Arndt & Reder, 2003). An especially powerful formatting manipulation involves presenting pictures versus printed words, with pictures drawing processing away from meaning and toward the surface appearance of the focal objects in the pictures (Israel & Schacter, 1997; Schacter, Israel, & Racine, 1999). The theoretical prediction about how such picture interference ought to affect age increases in false memory is straightforward: If younger children are less likely than older children or adults to process meaning relations when targets are presented as words, picture presentation will interfere less with such processing in younger children (because they do less of it in the first place), and hence, picture presentation will reduce false memory more in older than in younger children. Howe (2006) confirmed this pre-

diction in his developmental studies of false recall of categorized lists. As mentioned earlier, when such lists were presented as words, false recall of unrepresented category exemplars more than doubled (from .17 to .38) between the ages of 5 and 11. When such lists were presented as pictures, however, this age trend vanished. The reason was that (a) pictorial presentation had no effect on false recall in 5-year-olds but (b) it lowered the false recall of 7- and 11-year-olds to the level of 5-year-olds. Consistent with the necessity hypothesis, therefore, a formatting manipulation that interferes strongly with semantic processing reduced false recall more in older children than in younger ones, to the point where the intrusion rates of older children dropped to the level of younger children.

Howe (2008) subsequently replicated these results by using picture versus word presentation of DRM lists. In this research, 5-, 7-, and 11-year-olds listened to eight 12-word DRM lists and performed free recall following the presentation of each list. As the lists were read, the targets were also presented as printed words, as black-and-white line drawings, or as color photographs. The first two presentation formats had been used by Howe (2006), and they produced analogous findings. False recall of missing words increased with age (from .15 to .22 to .40) with the word format, but the false recall levels for black-and-white pictures (.10, .12, and .14) did not differ as a function of age. Note that, as before, pictorial presentation suppressed false recall more in older than in younger children, just as the necessity hypothesis predicts. However, Howe (2006) obtained another result that seemed quite counterintuitive. When the presentation format involved color photographs, false recall of missing words increased with age (from .12 to .16 to .40). Note that these values are virtually the same as the corresponding ones for words. Thus, for some reason, color photographs did not interfere with semantic processing of DRM lists, although black-and-white photographs did. Why? In a series of follow-up experiments, the answer was found to lie in differences between the background content of the two types of



pictures. With black-and-white drawings, there was a single, homogeneous background (the color white) for each depicted target, naturally. With color photographs, however, the backgrounds against which the depicted targets were shown were highly variable. Howe (2006) discovered that when background variability was eliminated and color photographs were used in which target objects were shown against a uniform background (e.g., a green screen), the expected formatting effect reappeared. That is, color photographs as well as black-and-white drawings eliminated the age increase in intrusions of missing words.

The third reduction manipulation is one that was also discussed earlier, in another context—namely, the presentation of child-generated versus adult-generated DRM lists. As mentioned, Metzger et al. (2008), Anastasi and Rhodes (in press), and Carneiro et al. (2007) all hypothesized that age increases in the DRM illusion are not due to the superior meaning connection abilities of older subjects but are due to the fact that young children do not grasp the meanings of some of the words on adult DRM lists. If so, they reasoned, age increases should disappear if subjects were exposed to child-generated DRM lists, but contrary to this prediction, the data showed that the DRM illusion increases with child- as well as adult-generated lists. However, note that this is also a necessity manipulation for which other predictions can be tested. To illustrate, consider the child- versus adult-generated lists for missing words such as *sleep*, *smell*, and *window*. For adult subjects, the adult-generated lists are the best exemplars of each of these concepts, whereas the child-generated lists, to the extent that they contain some different words, are not the best exemplars for adults. Thus, for adults, the child-generated lists for which the composition differs substantially from adult-generated lists interfere with the formation of meaning relations that support intrusions and false alarms. If so, the necessity hypothesis says that although age increases should be observed with both types of lists, the increases will be attenuated with child-generated lists because intrusions and false alarms will be selectively suppressed in older subjects. Age increases are predicted for both types of lists because even if words are selected so as to optimize children's comprehension of their meaning, adults will still be better at forming semantic relations between those words when they are presented together in a single list (Carneiro et al., 2007). This prediction can be evaluated with two of the articles because age was factorially crossed with list type. Metzger et al.'s false recognition data confirmed the prediction. Between the ages of 7 and 21, false recognition of missing words more than tripled (from .23 to .72) with adult-generated lists but only doubled (from .22 to .44) with child-generated lists. Metzger et al. did not observe larger age increases in false recall with adult-generated lists, and Anastasi and Rhodes did not observe larger age increases in either false recall or false recognition with adult-generated lists.

One cannot be sure, without further experimentation, why Metzger et al. (2008) and Anastasi and Rhodes (in press) obtained these null effects, but restricted range artifacts are obvious possibilities. Any design feature that restricts the level of false memory in adults for adult-normed lists will make it more difficult, of course, to detect larger age increases for such lists. In adult norms (Roediger et al., 2001; Stadler et al., 1999), the probability of false recall for the six adult-generated lists that Metzger et al. used (*car*, *doctor*, *foot*, *fruit*, *sleep*, *sweet*) is .44, but the probability of false recall in Metzger et al.'s adults was only .16—hence, the re-

stricted range problem. A likely reason for this low value is that Metzger et al. administered short DRM lists (eight words). In Anastasi and Rhodes's research, one-third of their adult-normed DRM lists were ones that produce the lowest levels of false memory in adults, and one-third were lists that produce only moderate levels of false memory in adults. To maximize the chances of detecting different rates of age change for child- versus adult-normed lists, it is necessary to use DRM lists that produce the highest levels of false memory in adults (because this maximizes the range of potential age variability; Brainerd et al., 2006).

### Conclusions

The results for manipulations of both the elimination and reduction variety are congruent with the hypothesis that, in connected-meaning tasks, older subjects' superior ability to form meaning relations across targets is a necessary precondition for the age trends that were reviewed earlier in this article. As expected under this hypothesis, age increases in false memory disappeared in conditions in which this superior ability could not be expressed (because meaning relations could not be formed among targets). Also as expected, age increases were attenuated or disappeared in conditions in which such meaning relations could be formed, but the method of target presentation made it difficult to do so.

When these conclusions are combined with the earlier ones for sufficiency manipulations, it can be seen that the extant literature on age increases in false memory establishes a one-to-one relation between such increases and developmental improvements in meaning connection. On the sufficiency side of this relation, manipulations that help children connect meaning across targets have been found to attenuate or eliminate age increases by elevating false memory in younger children. On the necessity side, manipulations that eliminate intertarget meaning relations or make it more difficult to process them have also been found to eliminate or attenuate age increases in false memory, but to do so by suppressing false memory in older children or adults.

### *Related Findings: Age Trends in Semantic Versus Nonsemantic Intrusions*

We now review a further set of findings that bear on the mechanism that FTT posits for age increases in false memories. These findings were not covered earlier because they do not involve either sufficiency or necessity manipulations. They are correlational results, and therefore, they provide weaker evidence than the findings that have already been reviewed. Nevertheless, they are probative because they involve direct measurements of age differences in meaning connection, and they tell a story that is consistent with the findings of sufficiency and necessity manipulations.

The results in question are concerned with age differences in semantic versus nonsemantic properties of the intrusions that children make during recall in connected-meaning tasks. As Mandler (1962) pointed out many years ago, data that bear on how information is organized in memory can be obtained by analyzing how subjects retrieve previously studied targets during free recall, with a central question being the extent to which output is organized along semantic versus nonsemantic lines. Traditionally, such data are generated by analyzing free-recall protocols for the degree to

which semantic and nonsemantic properties of study lists affect the order in which targets are output during recall. Examples of nonsemantic relations among targets include similarity of presentation positions (e.g., early, middle, late) and phonological similarity (e.g., rhymes, such as *battle*, *cattle*, and *saddle*). Examples of semantic relations among targets include membership in taxonomic categories (e.g., *violin-guitar-trumpet*, *pig-sheep-cow*) and synonymy (e.g., *hill-mountain-peak*, *couch-davenport-sofa*). A classic finding about memory development is that young children's recall is organized more along nonsemantic lines and less along semantic lines than the recall of older children and adults (Bjorklund, 1987; Bjorklund & Muir, 1988). For instance, the recall of young children usually reflects targets' presentation positions (e.g., the last few targets tend to be output together at the start of recall), targets' phonological relations (e.g., rhymes are output together), and targets' associative relations (e.g., associatively-related targets are output together). On the other hand, unless procedures are used that help young children to process the meaning content of targets, their recall usually does not reflect semantic relations among targets (e.g., same-category exemplars or synonyms are not clustered together during output). By late childhood, however, the organization of output during recall begins to reflect semantic relations among targets, without the need of such semantic support (Bjorklund & Hock, 1982; Bjorklund & Jacobs, 1985).

Brainerd et al. (2002) reasoned that if age increases in false memory are due to developmental improvements in meaning making, a similar age pattern should be observed in semantic properties of intrusions. In a prior adult experiment by Payne et al. (1996), the authors classified the intrusions that subjects made during free recall of DRM lists as belonging to one of three categories: missing word (e.g., *doctor* when recalling that list), semantically-related words (e.g., *disease* or *scalpel* when recalling the *doctor* list), and unrelated words (e.g., *pirate* or *guitar* when recalling the *doctor* list). They reported that virtually all of the intrusions were either missing words or semantically-related words. In short,

adults "get the gist" of DRM lists because they seem to understand that words that do not have some sort of medical meaning could not possibly have been on the *doctor* list, that words that do not have some sort of musical meaning could not possibly have been on the *music* list, and so forth. Brainerd et al. (2002) hypothesized that this would not be true for young children—that young children would not get the gist of DRM lists as well as adults and, hence, unrelated words would intrude at higher rates. It will be remembered that two of their experiments dealt exclusively with DRM recall, while the third was a recall-plus-recognition design. Brainerd et al. (2002) classified the intrusions that subjects made as missing words, semantically-related words, or semantically-unrelated words. The averages for the three types of intrusions, pooled over the three experiments, can be seen in Figure 6, where missing words and semantically-related words have been combined into a single value (semantic intrusions). Consistent with their hypothesis and with the notion that age increases in false memory are rooted in developmental improvements in meaning connection, 5- and 7-year-olds did not seem to get the gist of DRM lists because semantically-unrelated intrusions were just as common as semantic intrusions. After age 7, however, the incidence of semantic intrusions rose sharply, and the incidence of nonsemantic intrusions dropped precipitously. The young adults displayed the same profile of virtually no nonsemantic intrusions that Payne et al. had reported.

The pattern of age increases in semantic intrusions and decreases in nonsemantic intrusions during recall of DRM lists has been replicated in subsequent experiments. For instance, Brainerd et al. (2006; Experiment 1) found that the percentages of 6-year-olds' free-recall protocols that contained semantic versus nonsemantic intrusions were equal (23% in each case), while the percentage of 11-year-olds' protocols that contained nonsemantic intrusions (3%) was much smaller than the percentage that contained semantic intrusions (34%). These authors also found that in learning-disabled children, the incidence of nonsemantic intrusions was higher and the incidence of semantic intrusions was lower

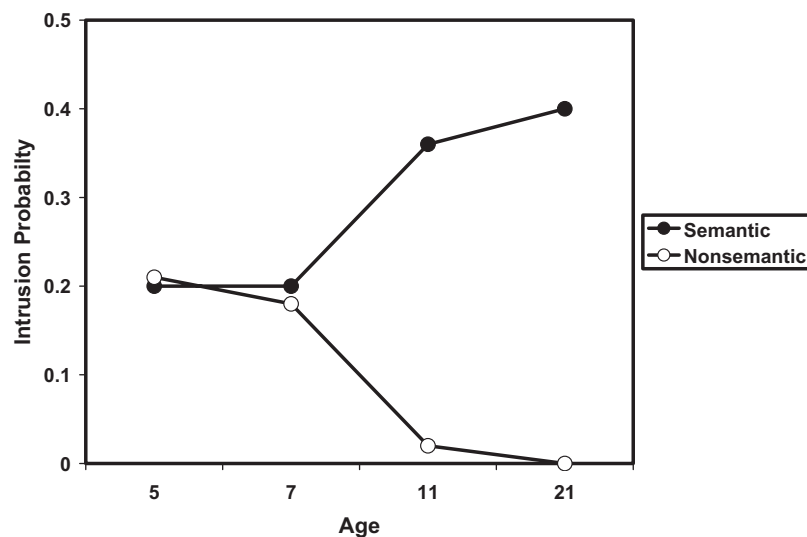


Figure 6. Semantic versus nonsemantic intrusions (age in years) in Experiments 1–3 of Brainerd, Reyna, and Forrester (2002).

than in same-age nondisabled children. In addition, Metzger et al. (2008) reported an age increase in semantically-related intrusions and a decrease in unrelated intrusions. Finally, Holliday et al. (2008) found that semantically-related intrusions increased with age, but they did not find that unrelated intrusions decreased with age. The latter finding was due to floor effects in unrelated intrusions at all age levels.

The last developmental study of age trends in semantic versus nonsemantic intrusions is especially informative because the authors attempted to determine whether young children's nonsemantic intrusions were utterly idiosyncratic or whether they reflected nonsemantic properties of targets. These findings were reported by Dewhurst and Robinson (2004). In their study, it will be remembered, 5-, 8-, and 11-year-old children listened to five 8-word DRM lists and performed free recall following the presentation of each list. Each intrusion that children made was classified as semantically related or semantically unrelated to the list that had just been presented. As in Brainerd et al.'s (2002, 2006) experiments, the mean number of protocols (out of a possible high of five) that exhibited semantic intrusions increased with age, and the mean number that exhibited nonsemantic intrusions decreased with age. More important, however, Dewhurst and Robinson classified nonsemantic intrusions as being of two types: unrelated to the presented list in any discernible way and phonologically related to that list in that they rhymed with one of the targets. It was found that phonological intrusions accounted for the great bulk of the nonsemantic intrusions and that, as a result, they were responsible for the age declines in nonsemantic intrusions. The overall age patterns for semantic and phonological intrusions in Dewhurst and Robinson's study are shown in Figure 7. There are two important conclusions. First, although young children's intrusions may be dominated by nonsemantic errors, those errors are not utterly unlawful and idiosyncratic. Rather, they may reflect properties of targets that young children are more adept at processing (e.g., phonological relations). Second, note the obvious convergence between the age declines in phonological intrusions in

Figure 5 with Holliday and Weekes' (2006) and Brainerd and Reyna's (2007) findings that false recall and false recognition of related distractors that sound like targets decline with age between early childhood and middle adolescence.

### Associative Activation

We have relied throughout on FTT's account of developmental reversals, both because it predicted such effects at a time when the opposite age trend was preeminent but also because convergent support for its assumptions has been supplied by a variety of necessity and sufficiency manipulations. An alternate hypothesis, called associative activation, has recently been proposed by Howe (2005, 2006) that deploys somewhat different ideas to explain reversal findings for two word-list tasks: the DRM illusion and categorized lists. Howe's hypothesis is derived from an early adult model of the DRM illusion, which assumes that errors are caused by associative rather than semantic processing—specifically, by backward associations that exist between missing words and list words (e.g., Deese, 1959). Backward associations refer to the fact that on tests of word association, some DRM list words elicit the missing word as a free associate. These backward associations are assumed to foment the illusion because, for instance, studying words such as *door* and *glass* causes *window* to be activated, to some extent, even though it is not presented. The primary evidence that has been adduced in favor of this assumption is that list differences in backward associative strength correlate positively with list differences in the strength of the illusion (Deese, 1959; Roediger et al., 2001).

Ghetti et al. (2002) and Brainerd et al. (2002) rejected the idea that associative activation predicts developmental reversals in the DRM illusion and concluded that, on the contrary, it predicts the standard developmental decrease pattern. Both groups of authors cited the classic developmental literature on word association, which shows that word association "is an early form of semantic relation that changes little in development" (Bjorklund & Jacobs,

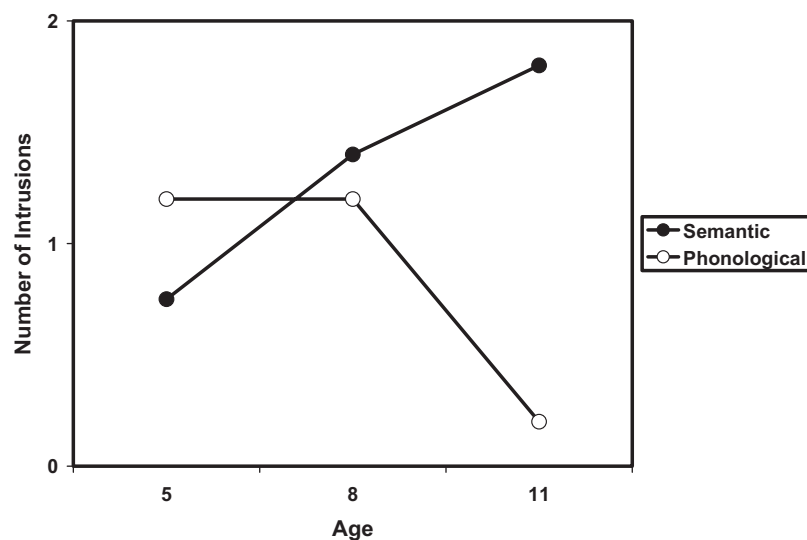


Figure 7. Semantic versus phonological intrusions at three age levels (in years) in an experiment reported by Dewhurst and Robinson (2004).

1985, p. 599) and that, therefore, presumably cannot be driving major age increases in false memory. Nevertheless, Howe (2005, 2006) argued that associative activation can be modified to accommodate reversals—not to predict them because the data are already in—by adding assumptions to the effect that two other aspects of associative processing increase with age: “the number and strength of inter-item associations in memory” (Howe, 2005, p. 1120) and “the automaticity with which inter-item associations are activated” (Howe, 2005, p. 1121). In this conception, studying *door* or *glass* causes *window* to be activated at all age levels, but activation is richer and more automatic in older subjects, and this is assumed to accentuate memory distortion. In support of this account, Howe (2005) has reported that variations in backward associative strength are positively correlated with intrusion rates for DRM lists and categorized lists, in children as well as adults.

As Howe’s (2005, 2006) account is very recent, a literature that evaluates its assumptions has yet to accumulate, and hence, it is premature to speculate as to how it will fare in experimental comparisons with theories such as FTT. However, it is clear that associative activation faces some major conceptual and empirical challenges that need to be overcome before comprehensive comparisons can be conducted. We note four examples.

The first is limited scope. The goal of any account of developmental reversals in false memory is to explain age increases for all the tasks that display this pattern. Most semantic memory illusions involve neither word lists nor word associations (cf. Brainerd & Reyna, 2005), whereas associative activation is confined to word-list tasks. The theory, as currently formulated, does not encompass the reversals that have been observed for Sloutsky and Fisher’s (2004b) categorized picture task, for the various suggestibility tasks that we reviewed, or for other connected-meaning tasks for which FTT predicts reversals (e.g., Brainerd & Mojardin, 1998) because word associations are not involved. The second challenge is that, as we saw earlier, explaining developmental reversals means explaining this pattern and simultaneously explaining the standard age-decrease pattern. In order to do that, a theory must posit opponent processes—specific mechanisms that stimulate errors (e.g., forming semantic relations between targets) and specific mechanisms that suppress them (e.g., verbatim-based recollection rejection). Otherwise, one cannot explain why some false-memory tasks, which seem very similar on the surface, exhibit opposite age trends. Although Howe’s account makes strong commitments to an error-stimulation process (associative rather than semantic processing), specific error-suppression mechanisms do not figure in the theory.

The third challenge is that empirical validation of the assumptions that are used to reconcile associative processing with development reversals is weak, those assumptions being (a) that there are developmental increases in the richness of word associations and in the automaticity of their activation and (b) that this somehow causes increased memory distortion. There are no data that we know of in which variations in the richness or automaticity of word associations have been causally tied to variations in children’s false memories, which would seem to be a prerequisite to comparative evaluations of the associative activation hypothesis. Against this view, it might be argued that because backward associations correlate with variations in children’s false memories, variations in the richness and automaticity of those associations must affect false memory. Beyond the circularity of such an

argument, the developmental literature is replete with dissociations between variables for which causal connections were once thought to be foregone conclusions, with dissociations between age increases in the ability to solve reasoning problems and the ability to remember problem information being well-worn illustrations (e.g., Bouwmeester et al., 2007). The fourth challenge is that the distinction between associative and semantic processing of word lists is a distinction without a difference because associative processing is semantic processing. As Cann et al. (2006) and Brainerd et al. (2007) have recently noted for DRM lists and as Anisfeld and Knapp (1968) noted long ago for word association generally, as backward associations between list words and missing words increase, so do familiar semantic relations such as antonymy, synonymy, functionality, emotionality, and so forth (see Table 2). Thus, if the backward association of *door* to *window* causes the latter to be processed, to some extent, when the former is studied, semantic relations (e.g., entity relations) are processed as well. The denouement, as has often been remarked in semantic-priming research, is that the confounding of associative with semantic processing means that no conclusions about associative versus semantic processing follow from data showing that word associations correlate with some aspect of memory performance.

Earlier in this section, we saw how semantic processing has been found to be both necessary and sufficient for certain forms of semantic false memory (and for developmental reversals). Therefore, it is important to ask: What is the corresponding relation between associative processing and semantic false memory? The answer, as Grossman and Eagle (1970) demonstrated some years ago, is that associative processing is sufficient but not necessary for semantic false memory. Associative processing obviously is not necessary because there are so many semantic false-memory tasks that involve neither word lists nor word associations. It is not necessary on word-list tasks either because, with paradigms such as the DRM illusion or categorized lists, significant levels of false memory are obtained for meaning-preserving distractors that are not associates of targets (e.g., Brainerd, Wright, Reyna, & Mojardin, 2001; Brainerd et al., 2004). However, associative processing is sufficient on these tasks because it stimulates specific forms of semantic processing that foment errors. Examples of such processing can be found in Tables 1 and 2.

### Concluding Remarks: The Next Steps

The research that has been reviewed supplies strong confirmation that false memories can increase markedly between early childhood and young adulthood. Although there may be other circumstances that foment such developmental reversals, it is now known that they occur when two, theoretically specified conditions are satisfied, conditions that also happen to be common in everyday life: (a) Memory errors are rooted in the formation of meaning connections between events, and (b) it is difficult to use verbatim memory for actual events to suppress those errors. Age increases in false memory have been detected with three methodologies that implement these conditions, which we shall refer to collectively as developmental-reversal paradigms. Several experiments with these paradigms included manipulations that were designed to establish causal connections between the two conditions, on the one hand, and age increases in false memory, on the other hand. Sufficiency manipulations (e.g., gist cuing) have been found to attenuate age



increases by elevating false memory more in younger subjects than in older ones, whereas necessity manipulations (both elimination and reduction manipulations) have been found to attenuate age increases by suppressing false memory more in older subjects than in younger ones.

Some core questions about developmental reversals in false memory have been asked and answered in the published research archive. However, such work is quite literally still in its infancy, the initial demonstration that the DRM illusion increases between early childhood and young adulthood having appeared only 5 years ago. Consequently, many fundamental questions remain to be answered, questions that will ultimately determine the theoretical and forensic yield of the developmental reversal effect. It is natural to ask what some of them might be, so that future research can build efficiently on what has been learned so far. As we surveyed this literature, some major unanswered questions became apparent, all of which would be productive targets for experimentation to aim at. We therefore conclude this article by considering some examples of questions for the next stage of research. They fall into two broad categories: (a) targets for basic research on the mechanisms that control age variability in developmental-reversal paradigms and (b) targets for applied research assessing the extent to which developmental reversals are relevant to the legal contexts that have supplied the historical motivation for the study of children's false memories.

#### *Some Objectives for Basic Research*

Of the various open questions about the processes that control age variability in developmental-reversal paradigms, the following four seem to be especially important targets: What role does the emergence of metacognitive processes play? Which brain regions are associated with age increases in false memory? Of the specific retrieval processes that have been linked to false memory in adults, which is most responsible for age increases in false memory? Why do pictorial materials exert such a powerful suppressive influence on age increases in false memory? We discuss these questions separately.

#### *The Role of Metacognition*

Given the centrality of metacognition in the larger literature on memory development, the lack of attention to metacognitive processes is a surprising feature of the literature on developmental reversals in false memory. Moreover, two considerations argue that such processes may contribute to the developmental reversal pattern. The first is the prominent role that metacognition has already played in the study of false memory in adults. Some of the earliest work in this vein was reported by Koriat and Goldsmith (1994, 1996), who introduced a technique called optional reporting to demonstrate that adults' knowledge of the logical properties and subjective qualities of true memories allows them to manipulate their output criteria on recall and recognition tests. They found that adults are able to fine tune their output criteria so as to minimize false alarms and intrusions of related distractors without reducing hits and correct recall of targets. Likewise, Dodhia and Metcalfe (1999) and Schacter et al. (1999) reported experiments that linked variability in the DRM illusion to specific metacognitive processes. Here, the distinctiveness heuristic of Schacter and associ-

ates (Dodson & Schacter, 2001, 2002a, 2002b; Schacter et al., 1999) has been an especially productive line of investigation. The second consideration is that developmental studies of metacognition show that such processes emerge during late childhood and early adolescence, which is the same age range when age increases in false memory are most pronounced. Although spontaneous meaning connection also improves during this age range and has already been linked to age increases in false memory, this does not preclude a causal role for metacognition as well. Indeed, some researchers have advanced this hypothesis (e.g., Ghetti & Alexander, 2004).

Yet, the studies that we have reviewed contain virtually no data on whether or how metacognitive processes contribute to developmental reversals in false memory. To fill that lacuna, a simple first step would be to determine whether children's performance on standard measures of metacognitive development (e.g., cf. Bjorklund, 2004) predicts their performance on false-memory tasks that exhibit developmental reversals. If the results are positive, the next logical step would be to determine whether causal links can be established for specific metacognitive processes, by manipulating children's reliance on those processes via the sorts of necessity and sufficiency designs that have been used to study the effects of meaning connection and verbatim processing. For example, two metacognitive processes that have already been investigated in adults, the distinctiveness heuristic of Schacter and associates (e.g., Schacter et al., 1999) and the criterion-tuning strategy of Koriat and Goldsmith (e.g., 1996), could easily be studied in necessity and sufficiency designs. In sufficiency designs, children of different age levels would be trained in how to use these strategies, whereas in necessity designs, manipulations would be imposed that interfere with the use of these strategies. In either case, the question would be whether age increases in false memory are attenuated.

#### *Brain Bases for Developmental Reversals*

The brain regions that are implicated in false memory have been topics of vigorous investigation in the adult literature. For connected-meaning illusions, in particular, several articles have appeared in which imaging techniques (e.g., functional magnetic resonance imaging), electrophysiological techniques (e.g., electroencephalography), and evidence from brain-damaged patients have been used to triangulate on the brain regions that are differentially activated by such tasks (for a review, see Schacter & Slotnick, 2004). Studies of this ilk have been reported with the DRM illusion (e.g., Cabeza, Rao, Wagner, Mayer, & Schacter, 2001) and with false memory for categorized materials (e.g., Garoff, Slotnick, & Schacter, 2005). Owing to such research, it is possible to formulate hypotheses about specific brain regions whose development may be tied to age increases in false memory.

In that connection, some key adult results are concerned with brain regions that are associated with extracting the gist of connected-meaning tasks and with the regions that are associated with the verbatim processing that suppresses errors on such tasks. In their literature review, Schacter and Slotnick (2004) concluded that neuroimaging studies of normal adults, together with behavioral studies of amnesic and Alzheimer's patients, point to the medial temporal lobes as a region of gist extraction. Other brain regions appear to be implicated in verbatim suppressive processes.

Here, Schacter and Slotnick concluded that neuroimaging data, electrophysiological data, and evidence from patient populations converge on the prefrontal cortex as a region of such processing. What do these findings suggest about developmental reversals in false memory?

Such evidence, together with the opponent-processes conception of false memory, generates predictions about what developmental neuroscience studies ought to reveal about the brain regions that are associated with such reversals. The same methodologies that produced converging evidence with adults have also been used with children (see Durston & Konrad, 2007; Henderson & Wachs, 2007; Karatekin, 2007; Luciana, 2007; Pennington, Snyder, & Roberts, 2007; Richmond & Nelson, 2007). If the brain regions that have been implicated in adult false memory are also important for age increases in false memory, a straightforward prediction would be that individual differences in maturation of the medial temporal lobes will be positively correlated with individual differences in developmental-reversal paradigms. Another straightforward prediction would be that within individual age groups, individual differences in maturation of the prefrontal cortex will be negatively correlated with individual differences in those paradigms. Further predictions are concerned with age differences in the patterns of brain activity in developmental-reversal paradigms. Remember, here, that young children show some evidence of susceptibility to connected-meaning illusions, and therefore, the question is whether the same brain regions are responsible for such illusions at older age levels, where susceptibility is much higher. If so, the expectation would be that the adult finding of differential activity in the medial temporal lobes and prefrontal cortices on such tasks would also be obtained in young children, although in reduced form. On the other hand, it is possible that increased susceptibility to connected-meaning illusions is due to a shift in the brain regions that are associated with such tasks. If so, the regions that display differential activity in developmental-reversal paradigms might be different in young children versus adolescents and adults. Some brain-imaging work that is designed to answer such questions is currently underway (e.g., Goodman, 2007; Paz-Alonso, Ghetti, Donohue, Goodman, & Bunge, in press).

### *Retrieval Processes That Control Developmental Reversals*

A central theoretical question in the adult memory literature, one that has not yet penetrated research on children, concerns the specific retrieval processes that underlie true- and false-memory responses. With respect to true memory, since the appearance of an early article by Strong (1913), several lines of evidence have pointed to the conclusion that two distinct retrieval processes are used to judge targets as old on recognition tests or to output them on recall tests, which are usually called recollection and familiarity (for a review of the modern literature, see Yonelinas, 2002). In recollective retrieval, subjects access vivid, realistic details of targets' (e.g., words') prior presentations, such as seeing visual details (e.g., font style, uppercase or lowercase letters) in the mind's eye or hearing auditory details (e.g., accent and gender of a voice) in the mind's ear. In the other form of retrieval, targets do not provoke realistic mental reinstatement of their prior presentation, but something about them seems so familiar that subjects are nevertheless confident that these items must have been presented.

Several methodologies have been used to measure these two forms of retrieval (cf. Yonelinas', 2002, review), but Tulving's (1985) remember/know procedure is the most frequently used technique. The consistent finding has been that recollection predominates over familiarity as the basis for recognizing or recalling targets (e.g., Gardiner & Java, 1991), with remember/know data suggesting that two-thirds to three-quarters of the hits on recognition tests are recollection-based (Donaldson, 1996).

What about false memories of unrepresented items? Because such items are not presented, realistic details of their prior presentation cannot be accessed, and common sense therefore argues that familiarity is the default operation that causes false memories. The first half of this statement is true by definition, but the second half does not necessarily follow. Although it is impossible to recollect realistic details of actual presentations, it is conceivable that subjects could recollect illusory realistic details of phantom presentations (i.e., subjects could see prior visual "presentations" of false-memory items flashing in the mind's eye and/or hear prior auditory "presentations" echoing in the mind's ear; Payne et al., 1996; Reyna et al., 2002). For instance, Reyna (1998) proposed that this could happen on recognition or recall tests if the strong gist memories that are created by connected-meaning tasks integrate fragmentary surface information, to form ersatz verbatim traces of unrepresented items. Because these vivid "presentation" details are necessarily illusory, this form of retrieval has been called phantom recollection in order to distinguish it from its true-memory counterpart.

If, as commonsense supposes, familiarity is the default retrieval operation in false memory, the finding that recollection predominates over familiarity for targets should be reversed for false memories. Many studies using the remember/know procedure have confirmed this prediction, using paradigms as varied as false memory for autobiographical experiences, photographs, and word lists (e.g., Conway, Collins, Gathercole, & Anderson, 1996; Heaps & Nash, 2001; Macrae, Schloerscheidt, Bodenhausen, & Milne, 2002). However, studies of adults' performance in developmental-reversal paradigms have yielded a different pattern. When remember/know tests are administered with the DRM illusion, for example, phantom recollection, not familiarity, predominates as the basis for this illusion. Roediger and McDermott (1995) reported that over three-quarters of false alarms to missing words involved vivid recollection rather than vague familiarity, and analogous results have been reported by several other investigators, using remember/know and other procedures (e.g., Brainerd, Payne, Wright, & Reyna, 2003; Lampinen, Meier, Arnal, & Leding, 2005; Payne et al., 1996). Likewise, with the second developmental-reversal paradigm, when adults study several exemplars of a category and then falsely recognize unrepresented exemplars, high levels of phantom recollection are detected with the remember/know procedure (e.g., Dewhurst, 2001; Dewhurst & Farrand, 2004). With the third developmental-reversal paradigm, when adults receive memory suggestions and then falsely recognize those suggestions as having been part of a previous experience, remember/know data show that phantom recollection can predominate over familiarity (e.g., Frost, 2000). Other connected-meaning illusions that have been studied with adults but not developmentally exhibit similar patterns. For instance, Reyna (2000) reported that phantom recollection predominated on tasks that measured false memory for unstated narrative events, when those events had been implied by several statements, and similar data

have been reported by Singer (2007; Singer & Remillard, in press). Recent research by Lampinen and associates (e.g., Lampinen et al., 2005) suggests that phantom recollections often consist of “borrowing” realistic presentation-phase details from related targets. With false memory for categorized materials, for example, subjects who falsely recognize unpresented exemplars of the color category, such as *green* or *purple*, might clearly recollect hearing these words pronounced in a male voice with a British accent, if such a voice had been used to present the actual exemplars. Importantly, such “borrowing” is driven by meaning connection because it does not occur for items that are inconsistent with the meaning of targets.

The relative contributions of different retrieval processes are just as important to the study of age increases in false memory as they are to the study of false memory in adults. A fundamental question is whether the proportionate contributions of phantom recollection and familiarity remain roughly constant or vary with development. Brainerd et al. (2004) pointed out that according to the traditional view that early childhood is a period of vivid imagination during which children have difficulty distinguishing fantasy from reality (e.g., the familiar phenomenon of believing that imaginary playmates are real; see also, developmental research on source monitoring, e.g., Foley & Johnson, 1985; Foley, Johnson, & Raye, 1983), the contribution of phantom recollection to false memories should decline with age and that of familiarity should increase. However, the opposite may be true, given older children’s spontaneous tendency to encode more features (semantic and perceptual) of a stimulus (Ceci, Lea, & Ringstrom, 1980), which would provide a broader basis for phantom recollection. Although age changes in the mix of different retrieval operations is an important question, perhaps the most significant motivation for studying these retrieval operations developmentally is that it provides an opportunity to discover where the counterintuitive process of phantom recollection comes from. As mentioned, it is known that connected-meaning illusions provoke high levels of phantom recollection in adults. Because these illusions increase markedly with age, it follows that absolute levels of phantom recollection must also increase with age, regardless of developmental trends in the relative mix of familiarity and phantom recollection. Why? What developmental changes are responsible for the growth of this illusory vivid retrieval process? Because phantom recollection is a hallmark of connected-meaning illusions in adults, a satisfactory account of age increases in false memory must answer these questions.

At present, however, the developmental literature contains virtually no evidence on the contributions of phantom recollection and familiarity to children’s false memories or on the contributions of recollection and familiarity to children’s true memories. The reason, as several authors have discussed (Brainerd et al., 1998; Ghetti et al., 2002), is methodological: The preeminent methodology for measuring these retrieval operations in adults, the remember/know procedure, is not appropriate for children because the instructions require 11-year-old levels of reading and language comprehension, and the length of the instructions imposes a high memory load. However, there are two other procedures that provide measurements of these operations, Jacoby’s (1991) process-dissociation paradigm and FTT’s conjoint-recognition paradigm (Brainerd, Reyna, & Mojardin, 1999), and that involve only simple, brief instructions that can be repeated as each probe is presented. The conjoint-recognition paradigm resembles the canonical

recognition procedure for measuring false memory that was described at the start of this article, except that subjects respond to recognition probes under three types of instructions: verbatim (accept only targets), gist (accept only related distractors), and verbatim + gist (accept both targets and related distractors). In the process-dissociation procedure, on the other hand, subjects are exposed to two distinct sets of target materials (List 1 and List 2), and they respond to recognition probes under two types of instructions: inclusion (accept both List 1 and List 2 targets) and exclusion (accept List 2 targets but reject List 1 targets). The applicability of these procedures to children has been investigated, and both have been found to produce reliable data in children as young as age 5 (Brainerd et al., 1998; Holliday & Hayes, 2000, 2001), and the conjoint-recognition procedure has also been found to produce reliable data in learning-disabled children (Brainerd & Reyna, 2002).

In short, there is no longer a methodological impediment to developmental research on the retrieval operations that control false memories. However, neither the process-dissociation procedure nor the conjoint-recognition procedure has yet been systematically applied in studies of developmental-reversal paradigms. We know of no published experiments of this sort with process dissociation and of only two with conjoint recognition. Brainerd et al. (2004) reported an experiment in which conjoint recognition was used to track the contributions of familiarity and phantom recollection to age increases in the DRM illusion, and they reported a second experiment in which conjoint recognition was used in the same manner with false memory for categorized materials. In the DRM experiment, three age levels were studied: 7, 11, and 14. It was found that age increases in this illusion were wholly attributable to age increases in phantom recollection (familiarity did not vary over this age range), and levels of phantom recollection at age 14 were well below the levels that have been reported in adult experiments. In the categorized materials experiment, two age levels were studied: 7 and 11. This experiment produced quite a different picture: Familiarity increased with age, while phantom recollection decreased. Thus, the literature does not yet contain even the beginnings of a clear picture of age changes in the retrieval processes that control false memory, but systematic application of the process-dissociation and conjoint-recognition procedures could remove this uncertainty.

### *Pictorial Presentation and Developmental Reversals*

The fourth objective for basic research is to clarify the unusually powerful influence that pictorial presentation exerts on age increases in false memory. Pictorial presentation has a special place in the history of memory development research. In many areas of investigation, particularly when the subjects are young children, it is traditional to present memory targets in pictorial form—as line drawings, as color photographs, as slide sequences, or as picture books that accompany stories—because they capture attention, are more concrete than equivalent verbal presentations, and because they ensure comprehension. This is also true of the study of children’s false memories, where pictorial presentation has been routinely used since the earliest studies (e.g., Ceci et al., 1987; Pezdek & Roe, 1995; Poole & Lindsay, 1995, 2002). As a practical matter, therefore, understanding how pictorial presentation modulates age changes is a topic of particular importance in any area of

memory development. Further, in false-memory research, we saw that the opponent-processes conception treats pictorial presentation as a prototypical verbatim-processing manipulation, that is, as a manipulation that produces stronger verbatim traces of targets and that can reduce false memories by making it easier for subjects to suppress related distractors (i.e., by retrieving verbatim representations of actual targets). For this reason, we saw that pictorial presentation has been used as a necessity manipulation in studies of developmental reversals in false memory.

Impressive experimentation along these lines has been reported by Howe (2005, 2008), which was summarized earlier. The results of this line of work were more dramatic than a mere attenuation of age increases. For children in the 5 to 11 age range, the usual age increases in both the DRM illusion and false memory for unrepresented category exemplars disappeared when DRM lists or categorized lists were presented as pictures rather than as words. The reason was that pictures drove false memory to near-floor levels throughout this age range. Gheiti et al. (2002) also failed to detect age increases in the DRM illusion with pictorial presentation, but this is not to say that pictorial presentation invariably eliminates age increases in false memory. It does not. In the studies of age increases in susceptibility to memory suggestion that we reviewed, pictorial presentation was used in the Pezdek and Roe (1995, 1997) and Ceci et al. (2007) experiments, and pictorial presentation was also used in Sloutsky and Fishers' (2004b; Fisher & Sloutsky, 2005) categorized materials paradigm. Also, when pictorial presentation of the DRM illusion is used with adults, although false-memory levels are lower than with word presentation, they are substantial in absolute terms (e.g., Israel & Schacter, 1997; Schacter et al., 1999). The latter finding, in particular, suggests that when pictorial presentation is used, there are age increases in false memory for both DRM lists and lists of category exemplars, but those age increases occur during middle and late adolescence, rather than during early and later childhood.

This still leaves the question of why pictorial presentation has such powerful suppressive effects on age increases in false memory. At least three hypotheses can be derived from the adult literature, none of which has been evaluated in developmental studies: (a) the distinctiveness heuristic; (b) enhanced verbatim processing; and (c) weakened gist processing. The first hypothesis grows out of Schacter and associates' studies of how picture versus word presentation affects the DRM illusion (e.g., Dodson & Schacter, 2002a, 2002b; Schacter et al., 1999). According to this line of work, adults possess metacognitive knowledge that pictures foment unusually vivid memories, so that on a memory test that follows exposure to a series of pictures, they expect that presented items will provoke such vivid memories, whereas unrepresented items will not. This is the distinctiveness heuristic. With respect to the DRM illusion, pictorial presentation of lists should activate this strategy but word presentation should not. The strategy will then suppress false memory for missing words in pictorial conditions because although, as we said, these items are often accompanied by illusory vivid "presentation" details, a substantial proportion are not, and those items will be weeded out by the distinctiveness heuristic. Note that because metacognitive knowledge increases with age, this would explain why pictorial suppression is more marked in older children than in younger children, but by itself, it would not explain why false memory following pictorial presentation apparently increases between mid-adolescence and young

adulthood (despite continuing increases in metacognitive knowledge). The second hypothesis, which grows out of the opponent-processes conception, is that pictorial presentation creates very strong verbatim traces of targets that are easily accessed on memory tests to suppress recall and recognition of related distractors. Because formation of the semantic relations that foment errors increases with age, this would also explain why pictorial suppression is more marked in older children (e.g., Reyna & Kiernan, 1995; Reyna & Lloyd, 1997; Reyna et al., 2007). The fact that false memory following pictorial suppression apparently increases between mid-adolescence and young adulthood is not problematic for this hypothesis because the ability to form semantic relations continues to increase during this age range. Finally, a third hypothesis, which also grows out of the opponent-processes conception, is that pictorial presentation is a semantic-interference manipulation; it impairs children's ability to connect meaning across different targets. Like the first and second hypotheses, the third one can handle the finding that pictorial suppression is more marked in older children because there is not much for pictures to interfere with in younger children. Like the second hypothesis, the third one can also handle the fact that false memory following pictorial presentation apparently increases between mid-adolescence and young adulthood because continuing improvements in the ability to form semantic relations reduce the effects of pictorial interference.

At present, all three of these hypotheses, or some combination of them, are viable candidates for explaining pictorial suppression of age increases in false memory. Differential evidence on the hypotheses would have direct implications for theoretical explanations of age increases in false memory. For instance, if the data favor either the second or third hypotheses, then pictorial suppression can be handled with current opponent-processes distinctions, thereby avoiding the introduction of additional theoretical concepts. However, if the data favor the first hypothesis, it would be necessary to enrich opponent-processes distinctions with metacognitive concepts.

### *Some Objectives for Applied Research*

A quarter century ago, the courts were struggling to understand the significance of the classic age-decline finding for false memory. Since then, owing to a generation of researchers who studied children's false memories in simulated forensic contexts, developmental data have come to figure centrally in cases that involve child witnesses and victims. Of course, children's heightened vulnerability to false memories has been of vital interest to the courts. In the adversarial arena of legal proceedings, attorneys introduce this finding, usually via expert scientific testimony, to achieve one of two objectives—to raise doubts about the reliability of evidence provided by individual children or, when conflicting evidence is offered by witnesses of different ages, as a basis for arguing that adults' (or older children's) evidence is inherently more reliable than younger children's. Depending on the case, data on the falsifiability of children's memories may be considered (a) during the pre-trial phase (e.g., prosecutors may decline to proceed with a case because they believe that children's evidence will be easily impeached by such data, or defense attorneys may cite such data in pre-trial motions to exclude children's testimony at trial); (b) during the trial phase (e.g., attorneys may introduce such data



to impeach the testimony of opposing witnesses or to buttress the testimony of their own witnesses); and (c) during the appeal phase (e.g., attorneys may argue that juries would not have convicted defendants if such data had been presented at trial). As mentioned earlier, the presentation of such data has become so commonplace that some courts have excluded expert testimony on children's vulnerability to false memories on the ground that it is common knowledge that is well understood by jurors (McAuliff et al., 2007).

Because it is now known that false memories also increase with age in experimental paradigms that implement the meaning-connectedness property of everyday remembering, the courts must ask to what extent should these new data influence scientific testimony about children's memories? The same question was posed 2 decades ago, in connection with the initial studies that showed age declines in false memory (e.g., see various chapters in Doris, 1991). The consensual response at the time was that the question required research in which children's false memories were studied in paradigms that simulate key features of legal cases. Since then, much experimentation of this sort has been reported (see various chapters in Toglia, Read, Ross, & Lindsay, 2007). For the courts to decide how the age-increase finding ought to influence scientific testimony, it is unarguable, we think, that the same type of forensically-guided experimentation must be undertaken. As starting points for such research, we discuss what we consider to be four lines of research whose forensic yield would be high: (a) age increases in false memory under conditions of emotional arousal; (b) age increases in false memory among disabled children; (c) age increases in false memory in eyewitness identification; and (d) age increases in susceptibility to memory suggestions about simulated forensic events. We consider these lines of research separately.

### *Emotional Arousal*

A hallmark of the events that children attempt to remember in legal cases is that they typically occurred in circumstances that were fraught with negative emotion. Some of the events are emotionally arousing in themselves (e.g., specific acts of physical violence that were committed during an assault), and others are not emotionally arousing in themselves (e.g., what the assailant was wearing) but may be influenced by the fact that children's emotions were aroused by other events. In either instance, levels of false memory may be affected by the presence of negative emotion, and the key question is how developmental variability in false memory is affected by such emotion.

Two basic methodologies for addressing this question are available in the adult literature, each of which simulates one of the ways in which emotion may affect memory for events in legal cases: recognition/recall of emotion-laden target materials and recognition/recall of neutral materials after an emotional state has been induced via other means. Concerning the first method, standardized pools of words and pictures are available (e.g., Bradley & Lang, 1999) that have been normed for emotional content. Such pools can be used to devise emotional analogues of developmental-reversal paradigms, that is, tasks in which target materials share meaning but are also emotion-laden. To illustrate, Budson et al. (2006) and Stein and Rohenkohl (2005) have developed DRM lists in which the missing words (e.g., *anger*) and the

corresponding targets (e.g., *mad*, *rage*, *annoyed*, . . . ) have negative emotional content. In the second method, a negative or positive mood is induced in subjects just before they participate in a connected-meaning illusion. This is most commonly accomplished by having positive-mood subjects listen to several minutes of fast, spritely music, while having negative-mood subjects listen to a comparable amount of slow, dirge-like music. For example, Storbeck and Clore (2005) and Corson and Verrier (2007) used this technique to induce positive and negative moods in their subjects, after which the subjects studied and recalled standard DRM lists.

Adult research with these paradigms has generally found that levels of true memory are higher and levels of false memory are lower in the presence of negative emotional states as compared with positive or neutral states. This suggests age increases in false memory may be attenuated, or may disappear altogether, in the presence of negative emotional states. Confirmation of this hypothesis would be of high forensic significance. Unfortunately, although both methodologies are readily applicable to children, virtually no developmental research with them has been published. We are aware of no published experiments in which the second methodology has been used with developmental-reversal paradigms, and only one in which the first methodology has been used. The latter study is the previously discussed experiment by Howe (2007), who administered Budson et al.'s (2006) emotional DRM lists to 5-, 7-, and 11-year-olds. Contrary to the hypothesis (from adult studies) that age increases in false memory will be attenuated by negative emotion, Howe's (2007) children exhibited higher levels of false memory for negative than for neutral lists, and age increases were not suppressed for negative relative to neutral lists. Considering that this is the only published study of its kind and that Howe's (2007) findings did not confirm expectations derived from adult research, much remains to be done when it comes to understanding how negative emotional states influence developmental-reversal paradigms.

### *Disability*

As research on developmental-reversal paradigms is broadened to encompass salient features of legal cases, it is essential to broaden subject samples to include children from populations with special characteristics that may influence susceptibility to false memories (e.g., Bruck, London, Landa, & Goodman, 2007). It is particularly important to include samples of children with learning disabilities, attention deficit, and retardation because, statistically speaking, the incidence of child witnesses and victims from these disability categories far exceeds their base rate in the population (e.g., Bruck & Ceci, 1999). Consequently, from the law's perspective, it is critical to determine which findings for children without disabilities extend to children in these disability categories. As a result, studies that include subject samples from these categories, especially samples of persons with retardation, have been regular features of the literature on memory suggestion (e.g., Gudjonsson & Henry, 2003; Henry & Gudjonsson, 2004, 2007). With tasks that produce the classic age-decrease pattern, the usual outcome has been higher levels of false memory in samples with learning disabilities and retardation than in control samples that have been matched on other variables. In addition to elevations in false memory, samples with learning disabilities and retardation usually display elevations in response bias.

With developmental-reversal paradigms, on the other hand, next to nothing is known about how disability may affect age increases in false memory. Only one experiment has been published in which a developmental-reversal paradigm was studied at different age levels with a subject sample that included children with learning disabilities and matched control children. This is the experiment by Brainerd et al. (2006) in which the DRM illusion was administered to groups of 7- and 11-year-olds with and without learning disabilities. The theoretical prediction was that age increases in false memory ought to be less marked in children with learning disabilities because they display reduced performance on measures of meaning connection, such as taxonomic clustering. The data were congruent with this prediction. In a related study, but one that lacked developmental comparisons, Weekes et al. (2008) found that the DRM illusion was weaker in a group of children with language disabilities than in same-age control children. In the only other related study of which we are aware, Carlin et al. (in press) compared levels of the DRM illusion in adults with retardation and a control sample of adults. Because adults with retardation also show reduced performance on measures of meaning connection, the theoretical prediction would be that they should display reduced levels of the DRM illusion. Instead, the illusion was more pronounced in the sample with retardation than in the control sample. However, interpretation of this pattern is complex because DRM lists were presented pictorially, which would be expected to interfere with the illusion in the control sample (e.g., Schacter et al., 1999) but not necessarily in the sample with retardation. Thus, as with the question of how emotional arousal modulates age increases in false memory, the question of how disability affects such increases remains entirely open.

### *Eyewitness Identification*

Eyewitness identification of suspects via photo spreads and lineups is the most prominent example of a memory test that supplies incriminating or exculpatory evidence in legal proceedings. It is also a measure that is known to be highly prone to false-memory responses (false-positive identifications) that arise from suggestion and other sources of distortion (Wells et al., 1998). It is so error prone that false-positive identifications are the leading cause of false convictions in the United States, and for that reason, the Department of Justice has promulgated a standardized field identification protocol that minimizes some known sources of distortion (Technical Working Group for Eyewitness Evidence, 1999). Developmentally, the accepted wisdom is that children are more susceptible to false-positive errors than adolescents or adults, and as previously noted, expert witnesses commonly opine on this point (Kassin, 2001). Although there are developmental studies that support that opinion, the literature is meager, and we reviewed an experiment by Ross et al. (2006) in which the susceptibility of eyewitness identifications to suggestion increased dramatically between early childhood and early adolescence, when culprits and innocent bystanders were of the same gender.

Beyond this, there are some common features of eyewitness identifications in the field that, in light of basic research on developmental-reversal paradigms, could cause false-positive rates to either increase or decrease, and developmental studies are urgently needed to decide between these alternatives. There are

some “facial gists” that are known to increase false positives, the standard ones being age, gender, ethnicity, and body build. When witnesses observe multiple people during the course of a crime, they are more apt to identify innocent people who match culprits in age or gender or ethnicity or build than people who do not match. Here, note the resemblance between the forensic situation of observing several people who share salient appearance features and the research situation of being exposed to several words or pictures that share salient meanings or perceptual features (e.g., phonology). As we have seen, research with developmental-reversal paradigms shows that semantic and perceptual relatedness produce opposing age changes in false memory, with semantic relatedness producing increases and perceptual relatedness producing decreases (Brainerd & Reyna, 2007; Holliday & Weekes, 2006). The aforementioned facial gists are definitely perceptual, because they involve the visual appearance of faces, but they are also semantic, because they can be defined and explained in words, and they tap sociocultural concepts.

Although these facial gists elevate adults' false-positive rates, it is not known whether this is due to semantic processing (e.g., a witness understands conceptually that because all of the people at a crime scene were young Hispanic men, young Hispanic men are the only possible culprits), or perceptual processing (e.g., it is more difficult to identify which of several faces is “correct” when they are all young Hispanic men than when they mismatch on age, ethnicity, or gender), or both. Research on this question in adults (e.g., Patterson & Baddeley, 1977) has shown that both could be involved because adults process both semantic and perceptual gist when they view faces. Without systematic developmental studies, therefore, it is unclear whether false-positive rates from matching of such facial gists will increase with age, because they depend on semantic processing, or decrease with age, because they depend on perceptual processing. It is even possible that the processing of some of these facial gists is predominantly semantic but the processing of others is predominantly perceptual, in which case developmental trends in error rates may be different for different facial gists.

### *Suggestibility of Eyewitness Memory*

It is fitting only to end with the paradigm that first focused attention on the classic age-decrease pattern in false memory, suggestibility of eyewitness memory. A standard criticism of the early experiments that established this pattern is that researchers had focused on memory suggestions about incidental details of innocuous experiences (e.g., what type of hat had been worn by a character in a story about leaving for school in the morning), rather than on suggestions about the sorts of events that are central in legal cases (see various chapters in Doris, 1991). In response to that criticism, many child-suggestibility experiments were conducted by using events that more closely simulate those that figure in certain types of legal cases, such as inappropriate touching (e.g., Pezdek & Roe, 1997); physical discomfort, pain, and claims of anogenital insertion (e.g., Bruck, Ceci, Francoeur, & Barr, 1995); property damage (e.g., Eisen & Carlson, 1998); genital examinations (e.g., Eisen et al., 2002); and behaviors that could be construed as sexualized (e.g., Poole & Lindsay, 1995).

It is not difficult to see that this same criticism can be raised with respect to the experiments that have demonstrated age in-

creases in suggestibility. Although those age increases were detected under theoretically-motivated conditions, three of the experiments (Ceci et al., 2007; Fazio & Marsh, in press; Pezdek & Roe, 1995) were concerned with suggestions about innocuous everyday experiences (e.g., play activities, baking a cake). Pezdek and Roe's (1997) experiment was the only one to produce age increases in suggestibility that dealt with forensically-relevant events (inappropriate touching) but that experiment was not specifically designed to search for age increases in suggestibility under connected-meaning conditions. Thus, the question of whether the suggestibility of eyewitnesses' memories for forensically-relevant events increases with age under such conditions, which would be of such high interest to the courts, remains entirely open.

### Afterword

Age trends in false memory have supplied test beds for counterintuitive theories of adult memory. The resulting findings—that false-memory illusions increase during childhood—have been surprising, challenging received wisdom about memory development and creating new dilemmas for the courts. Those findings have been widely replicated in independent laboratories in multiple countries. As is often the case with theoretically-motivated research, the findings raise more questions than they answer, questions about brain maturation, forensic applications, emotional arousal, and, indeed, fundamental mechanisms of false-memory facilitation and suppression (e.g., by pictures). Although the studies that have been reviewed represent a shift away from the traditional development-as-improvement view, Piaget (1968) anticipated developmental reversals in false memory based on the growth of understanding. The data suggest that understanding—meaning making—is the culmination of development and is therefore a better guide than verbatim accuracy in the study of memory and cognitive development.

### References

- Ackerman, B. P. (1992). The sources of children's source errors in judging causal inferences. *Journal of Experimental Child Psychology, 54*, 90–119.
- Ackerman, B. P. (1994). Children's source errors in referential communication. *Journal of Experimental Child Psychology, 58*, 432–464.
- Ackil, J. K., & Zaragoza, M. S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology, 60*, 57–83.
- Ackil, J. K., & Zaragoza, M. S. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology, 34*, 1358–1372.
- Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93*, 203–231.
- Anastasi, J. S., & Rhodes, M. G. (in press). Examining differences in the levels of false memories in children and adults using child-normed lists. *Developmental Psychology*.
- Anisfeld, M., & Knapp, M. (1968). Association, synonymity, and directionality in false recognition. *Journal of Experimental Psychology, 77*, 171–177.
- Arndt, J., & Gould, C. (2006). An examination of two-process theories of false recognition. *Memory, 14*, 814–833.
- Arndt, J., & Reder, L. M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory & Language, 48*, 1–15.
- Binet, A. (1900). *La suggestibilité*. Paris: Schleicher Feres.
- Bjorklund, D. F. (1987). How changes in knowledge base contribute to the development of children's memory. *Developmental Review, 7*, 93–130.
- Bjorklund, D. F. (2004). *Children's thinking: Cognitive development and individual differences*. Belmont, CA: Wadsworth.
- Bjorklund, D. F., Bjorklund, B. R., & Brown, R. D. (1998). Children's susceptibility to repeated questions: How misinformation changes children's answers and their minds. *Applied Developmental Science, 2*, 99–111.
- Bjorklund, D. F., Cassel, W. S., Bjorklund, B. R., Brown, R. D., Park, C. L., Ernst, K., & Owen, F. A. (2000). Social demand characteristics in children's and adults' eyewitness memory and suggestibility: The effect of different interviewers on free recall and recognition. *Applied Cognitive Psychology, 14*, 421–433.
- Bjorklund, D. F., & Hock, H. H. (1982). Age differences in the temporal locus of memory organization in children's recall. *Journal of Experimental Child Psychology, 33*, 347–362.
- Bjorklund, D. F., & Jacobs, J. W. (1985). Associative and categorical processes in children's memory: The role of automaticity in the development of free recall. *Journal of Experimental Child Psychology, 39*, 599–617.
- Bjorklund, D. F., Miller, P. H., Coyle, T. R., & Slawinski, J. L. (1997). Instructing children to use memory strategies: Evidence of utilization deficiencies in memory training studies. *Developmental Review, 17*, 411–441.
- Bjorklund, D. F., & Muir, J. E. (1988). Children's development of free recall memory: Remembering on their own. *Annals of Child Development, 5*, 79–123.
- Bouwmeester, S., Vermunt, J. K., & Sijtsma, K. (2007). Development and individual differences in transitive reasoning: A fuzzy trace theory approach. *Developmental Review, 27*, 41–74.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.
- Brainerd, C. J., Forrest, T. J., Karibian, D., & Reyna, V. F. (2006). Development of the false-memory illusion. *Developmental Psychology, 42*, 662–679.
- Brainerd, C. J., Holliday, R. E., & Reyna, V. F. (2004). Behavioral measurement of remembering phenomenologies: So simple a child can do it. *Child Development, 75*, 505–522.
- Brainerd, C. J., & Mojardin, A. H. (1998). Children's spontaneous false memories for narrative statements: Long-term persistence and retesting effects. *Child Development, 69*, 1361–1377.
- Brainerd, C. J., Payne, D. G., Wright, R., & Reyna, V. F. (2003). Phantom recall. *Journal of Memory & Language, 48*, 445–467.
- Brainerd, C. J., & Reyna, V. F. (1995). Autosuggestibility in memory development. *Cognitive Psychology, 28*, 65–101.
- Brainerd, C. J., & Reyna, V. F. (1996). Mere memory testing creates false memories in children. *Developmental Psychology, 32*, 467–476.
- Brainerd, C. J., & Reyna, V. F. (1998a). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology, 71*, 81–129.
- Brainerd, C. J., & Reyna, V. F. (1998b). When things that never happened are easier to remember than things that did. *Psychological Science, 9*, 484–489.
- Brainerd, C. J., & Reyna, V. F. (2002). Recollection rejection: How children edit their false memories. *Developmental Psychology, 38*, 156–172.
- Brainerd, C. J., & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review, 24*, 396–439.
- Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. New York: Oxford University Press.
- Brainerd, C. J., & Reyna, V. F. (2007). Explaining developmental reversals in false memory. *Psychological Science, 18*, 442–448.



- Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development, 73*, 1363–1377.
- Brainerd, C. J., Reyna, V. F., & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. *Journal of Memory & Language, 34*, 157–185.
- Brainerd, C. J., Reyna, V. F., & Mojardin, A. H. (1999). Conjoint recognition. *Psychological Review, 109*, 160–179.
- Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review, 110*, 762–784.
- Brainerd, C. J., Stein, L., & Reyna, V. F. (1998). On the development of conscious and unconscious memory. *Developmental Psychology, 34*, 342–357.
- Brainerd, C. J., Wright, R., Reyna, V. F., & Mojardin, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 307–327.
- Brainerd, C. J., Yang, Y., Howe, M. L., Reyna, V. F., & Mills, B. A. (2008). *The gist of the DRM illusion*. Manuscript submitted for publication.
- Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*, 331–380.
- Bruck, M., & Ceci, S. J. (1995). *Amicus* brief for the case of *State of New Jersey v. Michaels* presented by committee of concerned social-scientists. *Psychology, Public Policy, and Law, 1*, 272–322.
- Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology, 50*, 419–439.
- Bruck, M., Ceci, S. J., Francoeur, E., & Barr, R. (1995). I hardly cried when I got my shot: Influencing children's reports about a visit to their pediatrician. *Child Development, 66*, 193–208.
- Bruck, M., London, K., Landa, R., & Goodman, J. (2007). Autobiographical memory and suggestibility in children with autism spectrum disorder. *Development and Psychopathology, 19*, 73–95.
- Budson, A. E., Todman, R. W., Chong, H., Adams, E. H., Kensinger, E. A., Krangel, T. S., & Wright, C. I. (2006). False recognition of emotional word lists in aging and Alzheimer disease. *Cognitive and Behavioral Neurology, 19*, 71–78.
- Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory memory. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 4805–4810.
- Cann, D. R., McRae, K., & Katz, A. N. (2006, November). *Knowledge types underlying backward associative strength and false recall in the Deese-Roediger-McDermott paradigm*. Paper presented at the annual meeting of the Psychonomic Society, Houston, TX.
- Carlin, M. T., Toglia, M. P., Wakeford, Y., Jakway, A., Sullivan, K., & Hasel, L. (in press). Veridical and false pictorial memory in individuals with and without mental retardation. *Applied Cognitive Psychology*.
- Carneiro, P., Albuquerque, P., Fernandez, A., & Esteves, F. (2007). Analyzing false memories in children with associative lists specific for their age. *Child Development, 78*, 1171–1185.
- Ceci, S. J., & Bruck, M. (1993). The suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403–439.
- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom*. Washington, DC: American Psychological Association.
- Ceci, S. J., & Bruck, M. (1998). The ontogeny and durability of true and false memories: A fuzzy trace account. *Journal of Experimental Child Psychology, 71*, 165–169.
- Ceci, S. J., Caves, R., & Howe, M. J. A. (1981). Long-term retention of information that is incongruous with children's prior knowledge. *British Journal of Psychology, 72*, 443–450.
- Ceci, S. J., & Friedman, R. D. (2000). The suggestibility of children: Scientific research and legal implications. *Cornell Law Review, 86*, 34–108.
- Ceci, S. J., & Howe, M. J. A. (1978). Semantic knowledge as a determinant of developmental differences in recall. *Journal of Experimental Child Psychology, 26*, 230–245.
- Ceci, S. J., Lea, S. E. G., & Ringstrom, M. D. (1980). Coding processes in normal and learning-disabled children: Evidence for modality-specific pathways to the cognitive system. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 785–797.
- Ceci, S. J., Papierno, P. D., & Kulkofksy, S. (2007). Representational constraints on children's suggestibility. *Psychological Science, 18*, 503–509.
- Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility in children's memory: Psycholegal implications. *Journal of Experimental Psychology: General, 116*, 38–49.
- Chae, Y., & Ceci, S. J. (2005). Individual differences in children's recall and suggestibility: The effect of intelligence, temperament, and self-perceptions. *Applied Cognitive Psychology, 19*, 383–407.
- Chaing, W. C., & Chiu, C. H. (2006). Children's category structure and false memories. *Psychologia, 49*, 178–192.
- Chi, M. T. H., & Ceci, S. J. (1987). Content knowledge: Its representation and restructuring in memory development. In H. W. Reese & L. Lipsett (Eds.), *Advances in Child Development and Behavior, 20*, 91–146.
- The Committee on the Judiciary of the House of Representatives. (2006). *Federal rules of evidence*. Washington, DC: United States Government Printing Office.
- Connolly, D. A., & Price, H. L. (2006). Children's suggestibility for an instance of a repeated event versus a unique event: The effect of degree of association between variable details. *Journal of Experimental Child Psychology, 93*, 207–223.
- Conway, M. A., Collins, A. F., Gathercole, S. E., & Anderson, S. J. (1996). Recollections of true and false autobiographical memories. *Journal of Experimental Psychology: General, 125*, 69–95.
- Corson, Y., & Verrier, N. (2007). Emotions and false memory: Valence or arousal? *Psychological Science, 18*, 208–211.
- Deese, J. (1959). On the prediction of occurrence of certain verbal intrusions in free recall. *Journal of Experimental Psychology, 58*, 17–22.
- Dewhurst, S. A. (2001). Category repetition and false recognition: Effects of instance frequency and category size. *Journal of Memory & Language, 44*, 153–167.
- Dewhurst, S. A., & Farrand, P. (2004). Investigating the phenomenological characteristics of false recognition for categorized words. *European Journal of Cognitive Psychology, 16*, 403–416.
- Dewhurst, S. A., Pursglove, R. C., & Lewis, C. (2007). Story contexts increase susceptibility to the DRM illusion in 5-year-olds. *Developmental Science, 10*, 374–378.
- Dewhurst, S. A., & Robinson, C. A. (2004). False memories in children: Evidence for a shift from phonological to semantic associations. *Psychological Science, 15*, 782–786.
- Dodhia, R. M., & Metcalfe, J. (1999). False memories and source monitoring. *Cognitive Neuropsychology, 16*, 489–508.
- Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it": Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review, 8*, 155–166.
- Dodson, C. S., & Schacter, D. L. (2002a). Aging and strategic retrieval processes: Reducing false memories with a distinctiveness heuristic. *Psychology and Aging, 17*, 405–415.
- Dodson, C. S., & Schacter, D. L. (2002b). When false recognition meets metacognition: The distinctiveness heuristic. *Journal of Memory & Language, 46*, 782–803.
- Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition, 24*, 523–533.
- Doris, J. (Ed.). (1991). *The suggestibility of children's recollections*. Washington, DC: American Psychological Association.



- Durston, S., & Konrad, K. (2007). Integrating genetic, psychopharmacological and neuroimaging studies: A converging methods approach to understanding the neurobiology of ADHD. *Developmental Review, 27*, 374–395.
- Eisen, M. L., & Carlson, E. B. (1998). Individual differences in suggestibility: Examining the influence of dissociation, absorption, and a history of child abuse. *Applied Cognitive Psychology, 12*, S47–S61.
- Eisen, M. L., Qin, J. J., Goodman, G. S., & Davis, S. L. (2002). Memory and suggestibility in maltreated children: Age, stress arousal, dissociation, and psychopathology. *Journal of Experimental Child Psychology, 83*, 167–212.
- Esposito, N. J. (1975). Review of discrimination shift learning in young children. *Psychological Bulletin, 82*, 432–455.
- Fazio, L. K., & Marsh, E. J. (in press). Older, not younger, children learn more false facts from stories. *Cognition*.
- Fisher, A. V., & Sloutsky, V. M. (2005). When induction meets memory: Evidence for a gradual transition from similarity-based to category-based induction. *Child Development, 76*, 583–597.
- Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions: A developmental comparison. *Child Development, 56*, 1145–1155.
- Foley, M. A., Johnson, M. K., & Raye, C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development, 54*, 51–60.
- Frost, P. (2000). The quality of false memory over time: Is memory for misinformation “remembered” or “known”? *Psychonomic Bulletin & Review, 7*, 531–536.
- Gallo, D. A. (2004). Using recall to reduce false recognition: Diagnostic and disqualifying monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 120–128.
- Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks*. New York: Psychology Press.
- Gardiner, J. M., & Java, R. I. (1991). Forgetting in recognition memory with and without recollective experience. *Memory & Cognition, 18*, 617–623.
- Garoff, R. J., Slotnick, S. D., & Schacter, D. L. (2005). The neural origins of specific and general memory: The role of the fusiform cortex. *Neuropsychologia, 43*, 847–859.
- Ghetti, S., & Alexander, K. W. (2004). “If it happened, I would remember it”: Strategic use of event memorability in the rejection of false autobiographical events. *Child Development, 75*, 542–561.
- Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology, 38*, 705–718.
- Glaser, R. E., & Resnick, L. B. (1972). Instructional psychology. *Annual Review of Psychology, 23*, 207–276.
- Goodman, G. S. (2006). Children’s eyewitness memory: A modern history and contemporary commentary. *Journal of Social Issues, 62*(4), 811–832.
- Goodman, G. S. (2007). *False memory increases from childhood to adulthood: Why are children more accurate?* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Boston, MA.
- Goodman, G. S., Quas, J. A., Batterman-Faunce, J. M., Riddlesberger, M. M., & Kuhn, J. (1994). Predictors of accurate and inaccurate memories of traumatic events experienced in childhood. *Consciousness and Cognition, 3*, 269–294.
- Goodman, G. S., & Schaaf, J. M. (1997). Over a decade of research on children’s eyewitness testimony: What have we learned? Where do we go from here? *Applied Cognitive Psychology, 11*, S5–S20.
- Grossman, L., & Eagle, M. (1970). Synonymity, antonymity, and association, in false recognition responses. *Journal of Experimental Psychology, 83*, 244–248.
- Gudjonsson, G. H., & Henry, L. (2003). Child and adult witnesses with intellectual disability: The importance of suggestibility. *Legal and Criminological Psychology, 8*, 241–252.
- Heaps, C. M., & Nash, M. (2001). Comparing recollective experience in true and false autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 920–930.
- Hege, A. C. G., & Dodson, C. S. (2004). Why distinctive information reduces false memories: Evidence for both impoverished relational-encoding and distinctiveness heuristic accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 787–795.
- Henderson, H. A., & Wachs, T. D. (2007). Temperament theory and the study of cognition–emotion interactions. *Developmental Review, 27*, 396–427.
- Henry, L. A., & Gudjonsson, G. H. (2004). The effects of memory trace strength on eyewitness recall in children with and without intellectual disabilities. *Journal of Experimental Child Psychology, 89*, 53–71.
- Henry, L. A., & Gudjonsson, G. H. (2007). Individual and developmental differences in eyewitness recall and suggestibility in children with intellectual disabilities. *Applied Cognitive Psychology, 21*, 361–381.
- Hershkowitz, I., Fisher, S., Lamb, M. E., & Horowitz, D. (2007). Improving credibility assessment in child sexual abuse allegations: The role of the NICHD investigative interview protocol. *Child Abuse & Neglect, 31*, 99–110.
- Holliday, R. E., & Hayes, B. K. (2000). Dissociating automatic and intentional processes in children’s eyewitness memory. *Journal of Experimental Child Psychology, 75*, 1–42.
- Holliday, R. E., & Hayes, B. K. (2001). Automatic and intentional processes in children’s eyewitness suggestibility. *Cognitive Development, 16*, 617–636.
- Holliday, R. E., Reyna, V. F., & Brainerd, C. J. (2008). Recall of details never experienced: Effects of age, repetition, and semantic cues. *Cognitive Development, 23*, 67–78.
- Holliday, R. E., Reyna, V. F., & Hayes, B. K. (2002). Memory processes underlying misinformation effects in child witnesses. *Developmental Review, 22*, 37–77.
- Holliday, R. E., & Weekes, B. S. (2006). Dissociated developmental trajectories for semantic and phonological false memories. *Memory, 14*, 624–636.
- Horvath, F., & Meesig, R. (1996). The criminal investigation process and the role of forensic evidence: A review of empirical findings. *Journal of Forensic Sciences, 41*, 963–969.
- Horvath, F., & Meesig, R. (1998). A content analysis of textbooks on criminal investigation: An evaluative comparison to empirical research findings on the investigative process and the role of forensic evidence. *Journal of Forensic Sciences, 43*, 133–140.
- Howe, M. L. (2005). Children (but not adults) can inhibit false memories. *Psychological Science, 16*, 927–931.
- Howe, M. L. (2006). Developmentally invariant dissociations in children’s true and false memories: Not all relatedness is created equal. *Child Development, 77*, 1112–1123.
- Howe, M. L. (2007). Children’s emotional false memories. *Psychological Science, 18*, 856–860.
- Howe, M. L. (2008). Visual distinctiveness and the development of children’s false memories. *Child Development, 79*, 65–79.
- Howe, M. L., Cicchetti, D., Toth, S. L., & Cerrito, B. M. (2004). True and false memories in maltreated children. *Child Development, 75*, 1402–1417.
- Howe, M. L., Gagnon, N., & Thouas, L. (in press). Development of false memories in bilingual children and adults. *Journal of Memory & Language*.
- In the Matter of Riley, Austin, and Shelby Blanchard v. John Blanchard, Calif. Sup. Ct. CK40772/CK40775 (July 2001).
- Israel, L., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review, 4*, 577–581.

- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, *30*, 513–541.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.
- Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Developmental Review*, *27*, 283–348.
- Kassin, S. M. (2001). On the “general acceptance” of eyewitness testimony research: A new survey of experts. *American Psychologist*, *56*, 405–416.
- Kassin, S. M. (2005). The psychology of confessions: Does innocence put innocents at risk? *American Psychologist*, *60*, 215–228.
- Kimball, D. R., Smith, T. A., & Kahana, M. J. (2007). The fSAM model of false recall. *Psychological Review*, *114*, 954–993.
- Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory & Language*, *29*, 133–159.
- Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, *123*, 297–315.
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory. *Psychological Review*, *103*, 490–517.
- Koutstaal, W., Reddy, C., Jackson, E. M., Prince, S., Cendan, D. L., & Schacter, D. L. (2003). False recognition of abstract versus common objects in older and younger adults: Testing the semantic categorization account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 499–510.
- Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory & Language*, *37*, 555–583.
- Koutstaal, W., Schacter, D. L., & Brenner, C. (2001). Dual task demands and gist-based false recognition of pictures in younger and older adults. *Journal of Memory & Language*, *44*, 399–426.
- Lampinen, J. M., Leding, J. K., Reed, K. B., & Odegard, T. N. (2006). Global gist extraction in children and adults. *Memory*, *14*, 952–964.
- Lampinen, J. M., Meier, C. R., Arnal, J. D., & Leding, J. K. (2005). Compelling untruths: Content borrowing and vivid false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 954–963.
- Lewin, K. (1951). *Field theory in social science: Selected theoretical papers*. New York: Harper & Row.
- Loftus, E. F. (1975). Leading questions and eyewitness report. *Cognitive Psychology*, *7*, 560–572.
- Luciana, M. (2007). Developmental cognitive neuroscience. *Developmental Review*, *27*, 277–282.
- Lyon, T. D. (1995). False allegations and false denials in child sexual abuse. *Psychology, Public Policy, and Law*, *1*, 429–437.
- Lyon, T. D., & Saywitz, K. J. (2006). From post-mortem to preventive medicine: Next steps for research on child witnesses. *Journal of Social Issues*, *62*(4), 833–861.
- Macrae, C. N., Schloerscheidt, A. M., Bodenhausen, G. V., & Milne, A. B. (2002). Creating memory illusions: Expectancy-based processing and the generation of false memories. *Memory*, *10*, 63–80.
- Mandler, G. (1962). From association to structure. *Psychological Review*, *69*, 415–427.
- Marche, T. A. (1999). Memory strength affects reporting of misinformation. *Journal of Experimental Child Psychology*, *73*, 45–71.
- Marche, T. A., & Howe, M. L. (1995). Preschoolers report misinformation despite accurate memory. *Developmental Psychology*, *31*, 554–567.
- Marsh, E. J., Meade, M. L., & Roediger, H. L. (2003). Learning facts from fiction. *Journal of Memory & Language*, *49*, 519–536.
- McAuliff, B. D., Nicholson, E., & Ravanshenas, D. (2007, March). *Hypothetically speaking: Can expert testimony improve jurors' understanding of developmental differences in suggestibility?* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Boston, MA.
- McFarlane, F., Powell, M. B., & Dudgeon, P. (2002). An examination of the degree to which IQ, memory performance, socio-economic status and gender predict young children's suggestibility. *Legal and Criminological Psychology*, *7*, 227–239.
- McGough, L. S. (1993). *Child witnesses: Fragile voices in the American legal system*. New Haven, CT: Yale University Press.
- Metzger, R. L., Warren, A. R., Price, J. D., Reed, A. W., Shelton, J., & Williams, D. (2008). Do children “D/R-M” like adults? False memory production in children. *Developmental Psychology*, *44*, 169–181.
- Nelson, D. L., McGivney, V. M., Gene, N. R., & Janczura, G. A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological Review*, *105*, 299–324.
- Nelson, D. L., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review*, *99*, 322–348.
- Orbach, Y., & Lamb, M. E. (2000). Enhancing children's narratives in investigative interviews. *Child Abuse & Neglect*, *24*, 1631–1648.
- Orbach, Y., & Lamb, M. E. (2001). The relationship between within-interview contradictions and eliciting interviewer utterances. *Child Abuse & Neglect*, *25*, 323–333.
- Paris, S. G., & Carter, A. Y. (1973). Semantic and constructive aspects of sentence memory in children. *Developmental Psychology*, *9*, 109–113.
- Patterson, K. E., & Baddeley, A. D. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 406–417.
- Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory & Language*, *35*, 261–285.
- Paz-Alonso, P. M., Ghetti, S., Donohue, S. E., Goodman, G. S., & Bunge, S. A. (in press). Neurodevelopmental correlates of true and false recognition. *Cerebral Cortex*.
- Pennington, B. F., Snyder, K. A., & Roberts, R. J., Jr. (2007). Developmental cognitive neuroscience: Origins, issues, and prospects. *Developmental Review*, *27*, 426–441.
- Pezdek, K., & Roe, C. (1995). The effect of memory trace strength on suggestibility. *Journal of Experimental Child Psychology*, *60*, 116–128.
- Pezdek, K., & Roe, C. (1997). The suggestibility of children's memory for being touched: Planting, erasing, and changing memories. *Law and Human Behavior*, *21*, 95–106.
- Piaget, J. (1968). *Memory and identity*. Worcester, MA: Clark University Press.
- Pipe, M. E., Gee, S., Wilson, J. C., & Egerton, J. M. (1999). Children's recall 1 or 2 years after an event. *Developmental Psychology*, *35*, 781–789.
- Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children*. Washington, DC: American Psychological Association.
- Poole, D. A., & Lindsay, D. S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology*, *60*, 129–154.
- Poole, D. A., & Lindsay, D. S. (2002). Reducing child witnesses' false reports of misinformation from parents. *Journal of Experimental Child Psychology*, *81*, 117–140.
- Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, *27*, 975–986.
- Quas, J. A., Qin, J. J., Schaaf, J. M., & Goodman, G. S. (1997). Individual differences in children's and adults' suggestibility and false event memory. *Learning and Individual Differences*, *9*, 359–390.
- Reyna, V. F. (1998). Fuzzy-trace theory and false memory. In M. Intons-

- Peterson & D. Best (Eds), *Memory distortions and their prevention* (pp. 15–27). Mahwah, NJ: Erlbaum.
- Reyna, V. F. (2000, November). *Phantom recollection of narratives*. Paper presented at the annual meeting of the Psychonomic Society, New Orleans, LA.
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1–75.
- Reyna, V. F., & Brainerd, C. J. (1998). Fuzzy-trace theory and false memory: New frontiers. *Journal of Experimental Child Psychology, 71*, 194–209.
- Reyna, V. F., Holliday, R., & Marche, T. (2002). Explaining the development of false memories. *Developmental Review, 22*, 436–489.
- Reyna, V. F., & Kiernan, B. (1994). 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.
- Reyna, V. F., Kiernan, B. (1995). Children's memory and interpretation of psychological metaphors. *Metaphor and Symbolic Activity, 10*, 309–331.
- Reyna, V. F., & Lloyd, F. (1997). Theories of false memory in children and adults. *Learning and Individual Differences, 9*, 95–123.
- Reyna, V. F., Mills, B., Estrada, S., & Brainerd, C. J. (2007). False memory in children: Data, theory, and legal implications. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *Handbook of eyewitness psychology* (pp. 479–507). Mahwah, NJ: Erlbaum.
- Richmond, J., & Nelson, C. A. (2007). Accounting for change in declarative memory: A cognitive neuroscience perspective. *Developmental Review, 27*, 349–373.
- Robinson, K. J., & Roediger, H. L., III (1997). Associative processes in false recall and false recognition. *Psychological Science, 8*, 231–237.
- Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented on lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803–814.
- Roediger, H. L., III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*, 385–405.
- Ross, D. F., Marsil, D. F., Benton, T. R., Hoffman, R., Warren, A. R., Lindsay, R. C. L., & Metzger, R. (2006). Children's susceptibility to misidentifying a familiar bystander from a lineup: When younger is better. *Law and Human Behavior, 30*, 249–257.
- Rotello, C. M. (2001). Recall processes in recognition memory. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp. 183–221). San Diego, CA: Academic Press.
- Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory & Language, 40*, 1–24.
- Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. *Neuron, 44*, 149–160.
- Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition, 30*, 1054–1064.
- Seamon, J. G., Luo, C. R., Schlegel, S. E., Greene, S. E., & Goldenberg, A. B. (2000). False memory for categorized pictures and words: The category associates procedure for studying memory errors in children and adults. *Journal of Memory & Language, 42*, 120–146.
- Shiffrin, R. M. (2003). Modeling memory and perception. *Cognitive Science, 27*, 341–378.
- Shiffrin, R. M., & Steyvers, M. (1997). Model for recognition memory: REM: Retrieving effectively from memory. *Psychonomic Bulletin & Review, 4*, 145–166.
- Singer, M. (2007, July). *Impact of phrases versus words on text familiarity*. Paper presented at the joint meeting of the Experimental Psychology Society and the Psychonomic Society, Edinburgh, Scotland.
- Singer, M., & Remillard, G. (in press). Veridical and false memory for text: A multi-process analysis. *Journal of Memory & Language*.
- Sloutsky, V. M., & Fisher, A. V. (2004a). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133*, 166–188.
- Sloutsky, V. M., & Fisher, A. V. (2004b). When development and learning decrease memory. *Psychological Science, 15*, 553–558.
- Small, W. S. (1896). Suggestibility of children. *Pedagogical Seminary, 13*, 176.
- Smith, R. E., & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review, 5*, 710–715.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34–50.
- Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for words that create false memories. *Memory & Cognition, 27*, 494–500.
- State of Arizona v. Gregory Speers, Ariz. Sup. Ct. CR2000-00472 (Yuma, Ariz, 2007).
- State of New Jersey v. Michaels, NJ Supr. Ct. 136 N.J. 299, 642 A.2d 1372 (1994).
- Stein, L. M., & Rohenkohl, G. (2005, November). *Made in Brazil: Effect of emotion on memory illusion*. Paper presented at the annual meeting of the Psychonomic Society, Vancouver, British Columbia, Canada.
- Stern, W. (1910). Abstracts of lectures on the psychology of testimony and the study of individuality. *American Journal of Psychology, 21*, 270–282.
- Sternberg, K. J., Lamb, M. E., Esplin, P. W., Orbach, Y., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free-recall prompts in the course of forensic interviews. *Journal of Applied Psychology, 86*, 997–1005.
- Storbeck, J., & Clore, G. L. (2005). With sadness comes accuracy; with happiness, false memory: Mood and the false memory effect. *Psychological Science, 16*, 785–791.
- Strong, E. K. (1913). The effect of time-interval upon recognition memory. *Psychological Review, 20*, 339–371.
- Sugrue, K., & Hayne, H. (2006). False memories produced by children and adults in the DRM paradigm. *Applied Cognitive Psychology, 20*, 625–631.
- Swanson, H. L. (1991). Learning disabilities, distinctive encoding, and hemispheric resources: An information-processing perspective. In J. E. Obrzut & G. W. Hynd (Eds.), *Neuropsychological foundations of learning disabilities* (pp. 241–280). San Diego, CA: Academic Press.
- Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement*. Washington, DC: United States Department of Justice.
- Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory, 7*, 233–256.
- Toglia, M. P., Read, J. D., Ross, D. F., & Lindsay, R. C. L. (Eds.). (2007). *Handbook of eyewitness psychology*. Mahwah, NJ: Erlbaum.
- Toglia, M. P., Ross, D. F., Ceci, S. J., & Hembrooke, H. (1992). The suggestibility of children's memory: A social-psychological and cognitive interpretation. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *Development of long-term retention* (pp. 217–244). New York: Springer-Verlag.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26*, 1–12.
- Tussing, A. A., & Greene, R. L. (1999). Differential effects of repetition on true and false recognition. *Journal of Memory & Language, 40*, 520–533.
- Varendonck, J. (1911). Les témoignages d'enfants dans un procès retent [The testimony of children in a famous trial]. *Archives de Psychologie, 11*, 129–171.
- Weekes, B. S., Hamilton, S., Oakhill, J., & Holliday, R. E. (2008). False



- recollection in children with reading comprehension difficulties. *Cognition*, 106, 222–233.
- Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, 23, 603–648.
- Whipple, G. M. (1909). The observer as reporter: A survey of the psychology of testimony. *Psychological Bulletin*, 6, 153–170.
- Whipple, G. M. (1911). The psychology of testimony. *Psychological Bulletin*, 8, 307–309.
- Whipple, G. M. (1912). The psychology of testimony and report. *Psychological Bulletin*, 9, 264–269.
- Whipple, G. M. (1913). The psychology of testimony and report. *Psychological Bulletin*, 10, 264–268.
- Wilburn, C., & Feeney, A. (in press). Do development and learning really decrease memory? On similarity and category-based induction in adults and children. *Cognition*.
- Wu, L. L., & Barsalou, L. W. (2007). *Grounding concepts in perceptual simulation: I. Evidence from feature generation*. Manuscript submitted for publication.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, 46, 441–518.
- Young, K., Powell, M. B., & Dudgeon, P. (2003). Individual differences in children's suggestibility: A comparison between intellectually disabled and mainstream samples. *Personality and Individual Differences*, 35, 31–49.
- Zaragoza, M. S., Graham, J. R., Hall, N. N., Hirschman, R., & Ben-Porath, Y. S. (Eds.). (1995). *Memory suggestibility, and eyewitness testimony in children and adults*. Thousand Oaks, CA: Sage.

Received July 30, 2007

Revision received December 5, 2007

Accepted December 14, 2007 ■

### Correction to Reichenberg and Harvey (2007)

In the article “Neuropsychological Impairments in Schizophrenia: Integration of Performance-Based and Brain Imaging Findings,” by Abraham Reichenberg and Philip D. Harvey (*Psychological Bulletin*, 2007, Vol. 133, No. 5, pp. 833–858), on page 837, right column, first paragraph; in Table 1 (p. 835); and in Table 2 (p. 843), the word *perseverations* was misspelled as *preservations*. In addition, on page 846, left column, third paragraph, the last word in the sentence was incorrect. The correct word should be *hyperactivation*.

DOI: 10.1037/0033-2909.134.3.382