Published in final edited form as:

Dev Rev. 2012 September ; 32(3): 224–267. doi:10.1016/j.dr.2012.06.008.

Reliability of Children's Testimony in the Era of Developmental Reversals

C. J. Brainerd and V. F. Reyna

Department of Human Development, Cornell University.

Abstract

A hoary assumption of the law is that children are more prone to false-memory reports than adults, and hence, their testimony is less reliable than adults'. Since the 1980s, that assumption has been buttressed by numerous studies that detected declines in false memory between early childhood and young adulthood under controlled conditions. Fuzzy-trace theory predicted reversals of this standard developmental pattern in circumstances that are directly relevant to testimony because they involve using the gist of experience to remember events. That prediction has been investigated during the past decade, and a large number of experiments have been published in which false memories have indeed been found to increase between early childhood and young adulthood. Further, experimentation has tied age increases in false memory to improvements in children's memory for semantic gist. According to current scientific evidence, the principle that children's testimony is necessarily more infected with false memories than adults' and that, other things being equal, juries should regard adult's testimony as necessarily more faithful to actual events is untenable.

Keywords

children's testimony; law; false memory; fuzzy-trace theory; developmental reversals

Introduction

To say that the reliability of child witnesses' memories has been a controversial topic is an understatement of rather large proportions. Along with recovery of repressed memories (e.g., Loftus & Ketcham, 1994), false eyewitness identifications (e.g., Wells et al., 1998), and false confessions (e.g., Kassin & Kiechel, 1996), it has been one of the most contentious areas of psycho-legal research during the past quarter-century (Ceci & Bruck, 1995). To understand why, it is necessary to turn back the clock to the 1980s and consider two developments that first focused attention squarely on the memories of child witnesses. One was legal—namely, the withering away of child competence statutes that had excluded evidence from child witnesses on grounds of presumptive unreliability. The second was a clash between the traditional view of children's memories as being infected with faulty recollections and a revisionist view, according to which children's memories for certain types of criminal events are quite accurate.

© 2012 Elsevier Inc. All rights reserved.

Corresponding author: C. J. Brainerd, B-43 MVR Hall, Cornell University, Ithaca, NY 14853, Telephone: 607-254-1163, FAX: 607-255-9856, cb299@cornell.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Taking competence statutes first, the important background consideration here is that certain types of cases usually cannot be prosecuted without evidence from child witnesses. The classic examples are private crimes in which children are victims—crimes that may not leave reliable physical evidence—such as sexual or emotional abuse or neglect. Other examples are common home-based crimes for which children are often the only independent witnesses, such as spousal violence and the production of controlled substances.

Historically, in most states, evidence from children was routinely excluded on the ground that they are not competent witnesses because, among other things, the line between fantasy and reality is blurry in children (McGeough, 1993). Their presumed inaccurate memories, both in the sense of poor recollection of actual events and elevated recollection of events that did not happen, was a key factor in this view of child witnesses, one that seemed to derive ample support from early research on the accuracy of child witnesses (Binet, 1900; Small, 1896; Stern, 1910; Varendonck, 1911; Whipple, 1909).

Momentum for revisiting and modifying exclusionary statutes gathered in the 1970s, when most states made physicians mandated reporters of suspected child abuse and the United States Congress passed The Child Abuse Prevention and Treatment Act, which specified that in addition to physical abuse, sexual and emotional abuse qualified as child abuse. A federal office for collecting and reporting statistics on child abuse was established in the wake of that law, and data that were subsequently reported seemed to show an alarming rise in child sexual abuse from year to year. For instance, the proportion of abuse and neglect reports that involved sexual abuse doubled during the 1980s (Poole & Lamb, 1998). That circumstance focused attention on prosecution of such crimes, not only to punish the guilty but also to save vulnerable children from repeated victimization. Between the mid-1970s and the beginning of the 1990s, the number of child sexual abuse prosecutions nearly doubled nationwide. An important aid to such prosecutions was that legal barriers to evidence from child witnesses had been removed, making it possible for child victims to testify against defendants (Ceci & Bruck, 1995; McGough, 1993). Eventually, in 2006, Congress passed Federal Rules of Evidence 601 (The Committee on the Judiciary of the House of Representatives, 2006), which made statutory exclusion of evidence from children witnesses very difficult by specifying that juries should decide how much weight to assign to such evidence.

A consequence of mandating broad admissibility of child testimony is that the burden of proof falls squarely on the shoulders of children's memories in crimes that leave no reliable physical evidence and to which children are the only witnesses (Brainerd & Reyna, 2005). That raises an obvious question: What about the traditional notion, which seemed to be well supported by data, that children are highly unreliable witnesses—so unreliable that early researchers had claimed that children's evidence can only mislead jurors? This brings us to the second development, the clash between traditional and revisionist answers to that question.

Without putting too fine a point on it, a new interpretation was put forward, particularly among child advocacy professionals, that even young children's memories for certain types of events—especially ones involving their own bodies, such as those that figure in child sexual abuse—are quite accurate and are therefore exceptions to the rule that susceptibility to false memory is too high for children to be reliable witnesses. The emergence of this position during the 1980s is documented in a review by Ceci and Friedman (2000). As they discuss, proponents argued that when, as in abuse and neglect, crimes are traumatic events *that involve children's bodies* and that they have therefore *directly experienced* (rather than merely observed), their recollections of those events are accurate and quite resistant to falsification. In short, reports of such events, even by young children, are not apt to be either spontaneous false memories or false memories that arise from external suggestion. That, in

turn, led to the recommendation that suggestive and leading questions could be used to stimulate disclosures during investigations of child abuse allegations because such questions were unlikely to stimulate false allegations (Ceci & Freidman, 2000). Because young children typically volunteer few recollections of anything when asked in a general way about the events of their lives, the use of suggestive and leading questions became prevalent in abuse investigations (Brainerd & Reyna, 2005). Although this served the highly desirable goals of prosecuting abusers and preventing re-victimization of children, by increasing the ease of investigative documentation of abuse from victims, there was a weak link in the revisionist position: It was based on little in the way of new scientific findings, and such findings as there were are open to methodological challenges (Brainerd & Reyna, 2005).

Not surprisingly, child memory researchers remained skeptical of the revisionist position and proceeded to evaluate it under controlled experimental conditions. What followed was a series of experiments that were ultimately interpreted as demonstrating that young children are *less resistant* to suggestion-induced false memories than older children or adults (e.g., Ceci, Ross, & Toggia, 1986)—including false memories of quasi-sexual events involving their bodies (e.g., Poole & Lindsay, 1995) and of physically and emotionally painful events involving their bodies (e.g., Bruck, Ceci, Franconeur, & Barr, 1995). We say “ultimately interpreted” because, from the start, there was some disagreement about whether the data actually showed that young children were less resistant to suggestion-induced false memories than older children or adults (see various chapters in Doris, 1991). That disagreement was stimulated by the fact that although age declines in suggestibility were detected in early studies such as Ceci et al.’s, they were not detected in other early studies (e.g., Howe, 1991; Marin, Holmes, Guth, & Kovac, 1979). Moreover, failures to detect age declines in suggestibility are not confined to early studies, a 2001 experiment by Poole and Lindsay being a case in point. Those authors reported a carefully designed study of 3- to 8-year-olds’ susceptibility to memory suggestions provided by their parents about staged events (science demonstrations). Both recognition and recall tests were used to measure false memory. Both types of tests showed that parental suggestion created false memories at all age levels, but age trends were different for recognition and recall. False memories declined by roughly 50% between 3 and 8 on recognition tests but did not vary with age on recall tests. Beyond the controversy over now-you-see-it-now-you-don’t age trends, an often-voiced counterargument to all of these experiments is that even massively consistent evidence of age declines in suggestibility would not completely cut the ground from under the revisionist view because researchers cannot, for ethical reasons, create false memories of events that the law classifies as crimes.

Such objections were eventually swept aside by two events, one scientific and the other public. The scientific event was the publication by Ceci and Bruck (1993) of the first comprehensive review of the developmental literature on memory suggestibility. Those authors acknowledged the controversy over now-you-see-it-now-you-don’t age trends, and carefully reviewed the evidence on both sides. They concluded their review with a table that, to most readers’ eyes, revealed a strong signal embedded in the noise. The table listed all developmental suggestibility experiments that had been reported to date, appending a plus mark to those that had detected age declines and a minus mark to those that had detected no age trends. (No experiment had detected age increases.) The table spoke volumes: There were pluses appended to 83% of the experiments.

The public event was when some sexual abuse prosecutions involving young children (primarily preschoolers) received widespread media coverage, provoking public outcry (for a review of these cases, see Ceci and Bruck, 1995). In the public mind, one reason that those prosecutions appeared questionable was that the acts of child sexual abuse for which the defendants were tried and, in some instances, convicted were bizarre and improbable. For

instance, in *State of New Jersey v. Michaels* (1994) a preschool teacher was convicted of 115 counts of sexual abuse involving 20 child victims, but the children's allegations against her included such strange behavior as playing the piano in front of the classroom while nude. Ultimately, the conviction was reversed by the New Jersey Supreme Court, which concluded that such bizarre allegations may have been false memories that were created by suggestive questioning. In other cases that captured public attention (e.g., *State of California v. Buckley*, 1990), details of weird acts of child sexual abuse were contradicted by physical evidence presented at trial.

As the only evidence of criminal acts in these cases consisted of children's allegations, which were often obtained with suggestive questioning, the cases raised reliability concerns among child memory researchers. The result has been the production of an extensive literature on false memory during the preschool-to-young-adult age range and on variables that increase and decrease susceptibility to false memory, work that continues to the present day. Susceptibility to spontaneous false memories and to false memories that are implanted via suggestion have both been extensively studied, and multiple reviews of both strands of research are available (see 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), some of them written by contributors to this special issue. Although the scientific study of children's false memory has produced a data archive of enduring legal significance, our interest in this section lies specifically with what has been called massive evidence of a global developmental decline in such errors between early childhood and young adulthood. The evidence of age decline is quite broad-based inasmuch as this trend has been detected with a variety of false-memory paradigms. With respect to spontaneous false memory, declines have been reported in paradigms that range from false memory for unrepresented synonyms and same-category exemplars of words presented on lists (e.g., Brainerd & Reyna, 1996; Brainerd, Reyna, & Brandse, 1995; Brainerd, Reyna, & Kneer, 1995), to false memory for paraphrases of literal and metaphorical statements (Reyna & Kiernan, 1994, 1995), to false memory for source details (Ackerman, 1992, 1994; Ackil & Zaragoza, 1995), to false memory for real-world events during free and cued recall of live event sequences (Pipe, Gee, Wilson, & Egerton, 1999; Poole & White, 1991), to false memory for words during free and cued recall of word lists (Bjorklund & Muir, 1988), to false memory for numerical information on mathematics problems (Brainerd & Gordon, 1994). The evidence of age declines in susceptibility to suggestion-induced false memories is even more vast and includes age declines in implanted false memories of thefts (e.g., Bjorklund, Bjorklund, & Brown, 1998; Bjorklund et al., 2000), events that children knowingly confabulate (Ackil & Zaragoza, 1998), personal traumatic experiences (e.g., Goodman, Quas, Batterman-Faunce, Riddlesberger, & Kuhn, 1994), quasi-sexual events involving children's bodies (e.g., Poole & Lindsay, 1995, 1996), and physically and emotionally painful events involving children's bodies (e.g., Bruck et al., 1995). Considering that the same studies generally show that true memories (i.e., for events that actually happened) increase with age as false memories are decreasing, the overall pattern is one of steady developmental improvements in net accuracy.

Such research had direct forensic implications for the reliability of children's evidence. Obviously, it echoed the traditional view that the line between fantasy and reality is not as sharply drawn in children as it is in adults. At a more specific level, the picture was that, on the one hand, children's memories are not as utterly unreliable as the law had once assumed, but on the other hand, key items of evidence in witnesses' reports are increasingly likely to be false memories as witnesses become younger and younger. This leads to two principles that juries can apply in judging the credibility of evidence that is provided by children (Brainerd, Reyna, & Zember, 2011). First, when children are the *only* sources of evidence that bears *directly* on guilt (as is common in cases involving child abuse and neglect), due

weight must be given to the fact that there is an elevated risk that the evidence is tainted by false memories. Second, when children and adults are both sources of evidence that bears directly on guilt (as is common in certain domestic crimes and child custody disputes), due weight must be given to the fact that children's evidence is more apt to be tainted by false memories than adults'. For the past two decades, ideas such as these and the scientific findings that support them have been centerpieces of expert testimony in thousands of cases in which evidence from children was presented. The accumulated weight of such testimony, coupled with intense media scrutiny of certain cases, has led some courts to rule that knowledge of child witnesses' susceptibility to false memories now extends beyond the scientific and legal communities to include the lay public (Brainerd et al., 2008). It has thus become increasingly frequent for courts to rule that because children's heightened susceptibility to false memories is common-sense knowledge, expert scientific testimony is no longer needed to establish this fact in court, and juries can simply be instructed to consider it in weighing the credibility of children's evidence (McAuliff, Nicholson, & Ravanshenas, 2007).

Against this background, our ultimate concern in this article lies with the fact that these established ideas about child witnesses are now under challenge by false memory research that has appeared during the past decade. This may seem surprising considering that the number of studies that show age declines in false memory (coupled with age increases in true memory) is so extensive. As we discuss in the first section below, however, there is a clear theoretical basis for predicting that in some situations that are of definite forensic relevance, false memories will grow with age and net accuracy will decline. That basis falls out of a standard account of adults' false memories, fuzzy-trace theory (FTT; Reyna & Brainerd, 1995; Brainerd & Reyna, 2005), as well as prior research in which FTT was used to predict developmental reversals in *reasoning* (situations in which reasoning biases and illusions wax rather than wane between early childhood and young adulthood; Reyna & Ellis, 1994). As we discuss in the second section below, many developmental studies have appeared that not only confirm counterintuitive reversals in the developmental trajectory of false memory and net age declines in memory accuracy, but more important for the courts, evaluate the memory processes that cause such reversals. This new developmental reversal pattern, like the pattern that preceded it, has extensive forensic implications, the most obvious of which is that the default assumption that children's evidence is more infected by false memories than adults' is questioned by data that show clear exceptions to this rule.

Theoretical Reasons for Developmental Reversals

Before the current surge of research on developmental reversals in false memory, some of the contributors to this special issue noted that there were theoretical reasons for predicting that certain varieties of false memories could increase dramatically and net accuracy could decline with age (Brainerd & Reyna, 1998; Ceci & Bruck, 1998; Reyna & Kiernan, 1994). Moreover, those varieties were of forensic relevance because, like memory for crimes, they are rooted in understanding and retrieving memory for meaningful events and the thematic connections that exist between experienced events ("the gist"). At about the same time, Brainerd and Mojardin (1998) reported a confirmation of the developmental reversal prediction, using a narrative memory paradigm. They found that when children listened to short narratives that established simple meaning relations among target items (*The coffee is hotter than the tea. The tea is hotter than the cocoa. The cocoa is sweet.*), the probability of falsely remembering hearing *unpresented* sentences that preserved narrative gist (*The tea is cooler than the coffee*) increased from 7% to 25% to 26% to 30% in 6-, 8-, 11-, and 20-year-olds, respectively.

The false memories that crop up in legal evidence are what memory researchers call *semantic* false memories because, like this narrative example, they preserve the meaning of true information.¹ What are the grounds for supposing that such distortions will sometimes increase between early childhood and young adulthood? They can be found in the prevailing approach to explaining semantic false memories in adults, opponent-processes models (Brainerd & Reyna, 2005). Theories of this sort were first formulated because older ones were unable to handle the experimental dissociations that were observed between true and false memories, across various tasks and manipulations (Brainerd & Reyna, 1992, 1995; Brainerd et al., 1995; Reyna & Kiernan, 1994, 1995). The core assumption is that there are distinct memory processes that contribute in opposite ways to true and false memory—in particular, processes that simultaneously enhance true memory and suppress false memory, and vice versa. One such model, FTT, was used by Ceci and Bruck (1998) and Brainerd and Reyna (1998) to generate developmental reversal predictions.

According to FTT, subjects store dissociated verbatim and gist representations of experience (Reyna & Brainerd, 1995). Verbatim traces are representations of events' surface features, such as the shape, color, size, and texture of an object (e.g., a *Coke* can), and superficial features of the contexts in which they are encountered. Gist traces, on the other hand, are representations of events' senses, patterns, and meanings—most commonly in memory research, their semantic features (e.g., soda) and meaningful contextual features as well. Thus, while the information in verbatim traces captures surface qualities that can be directly experienced in events, the information in gist traces must be accessed using events as retrieval cues and is therefore more subject to individual differences in knowledge and learning history (Brainerd & Reyna, 1995). Because verbatim and gist traces are tagged with contextual features, they are *episodic* representations; personal records of the events of our lives, which are retrieved when we respond to memory tests that inquire about the putative content of those events. On such tests ("Did you drink a Coke at lunch?"), either verbatim traces, gist traces, or both may be accessed.

FTT posits that although both types of traces support true memory, they have opposite effects on false memory. If you drank a Coke at lunch, retrieving a verbatim trace of holding a red Coke can in the cafeteria or a gist trace of ordering a soda support true memory for that event because the surface features of the former and the semantic features of the latter both match the surface and semantic features of the putative event. When it comes to false memories of related events, such as drinking a Pepsi, Sprite, or 7-Up, retrieving this same verbatim trace suppresses them because surface features mismatch, but retrieving the corresponding gist trace supports them because semantic features match (Brainerd & Reyna, 2002). Hence, false memory for events that preserve salient aspects of the meaning content of experience depends on which types of traces are accessed, which in turn will depend on the sorts of retrieval cues that are supplied and on the relative availability of verbatim and gist traces in storage. The latter is especially subject to age variability and is a key consideration in predicting developmental reversals in false memory.

These few distinctions about the representations that foment versus suppress false memories suggest a general approach to increasing the strength of memory distortions, which can then be filled in with a range of specific manipulations. In that approach, levels of false memory ought to rise when a manipulation has either or both of two effects: (a) It increases subjects' tendency to access gist traces on memory tests, or (b) it reduces their tendency to access verbatim traces on memory tests. A manipulation could produce such effects on either the front or back end of an experiment; that is, the effects could be generated by influencing the tendency to store these distinct types of traces in the first place (an *availability* influence) or

¹By "semantic" memory, we mean memory based on meaning (rather than the store of memory for facts).

by influencing the tendency to retrieve them on memory tests (an *accessibility* influence) or both. Front- and back-end manipulations have both been investigated in the adult literature (for a review, see Brainerd & Reyna, 2005). A simple front-end manipulation that should strengthen gist traces relative to verbatim traces consists of exposing subjects to several distinct events that all share the same salient meaning. In our illustration of drinking a Coke, subjects might be exposed to a written narrative or a video in which a central character named Bob drinks one of several sweet, fizzy, sodas at different meals—say, 7-Up, A&W, Coke, Fresca, Mountain Dew, RC, Sprite, Squirt, and Vernors. This procedure repeatedly instantiates “soda,” presumably creating very strong gist traces of that meaning, but it does not create correspondingly strong verbatim traces of the individual drinks because each appeared only once in the story. Sometime later, subjects respond to recognition tests on which the false memory items ask whether Bob consumed other familiar sodas (e.g., Crush, Dr. Pepper, Jolt, Pepsi, and Tab). Relative to a control condition in which Bob only drinks a Coke, subjects should display elevated levels of false memory for these other sodas, for two reasons. First, the type of gist memory that supports such distortions has been greatly strengthened (and is thus more apt to be retrieved), in comparison to verbatim traces of the individual drinks (Reyna & Lloyd, 1997). Second, even if a verbatim trace (say, that Bob drank an RC) is retrieved when one of these false-memory items (say, that Bob drank a Pepsi) is tested, it will not be as effective at suppressing an error because subjects know that Bob drank many sodas other than RC, one of which may have been a Pepsi (Brainerd, Reyna, Wright, & Mojardin, 2003).

Here, we should remind ourselves of a point that is pertinent to our ultimate concern with the reliability of child witnesses’ memories: Procedures in which subjects experience many events that share salient meaning are analogues to everyday remembering. For instance, consider the numerous examples of meaning-sharing objects and events that populate our experience as we move through our daily activities (eating breakfast, driving to work, attending class, shopping for groceries, attending a sporting event). Procedures that echo this property of real-world experience—which, for obvious reasons, are called connected-meaning tasks—have often been found to elevate false memory. Indeed, a paradigm that has been intensively studied in the adult literature, the Deese/Roediger/McDermott (DRM; Deese, 1959; Roediger & McDermott, 1995) illusion, is just such a task. The DRM illusion involves studying short lists of 12–15 familiar words, followed by free recall tests or recognition tests. The lists are special ones that have been constructed from norms of word association (e.g., Nelson, McEvoy, & Schreiber, 1999). To generate such a list, a familiar word is identified that has many forward associates in such norms (e.g., words such *anger*, *chair*, *doctor*, and *sweet*), and the first 12–15 forward associates are selected (for *chair*, its first 15 forward associates are *table*, *sit*, *legs*, *seat*, *couch*, *desk*, *recliner*, *sofa*, *wood*, *cushion*, *swivel*, *stool*, *sitting*, *rocking*, *bench*). This is clearly a connected-meaning list owing to the all the familiar semantic relations that the words share. [A quantitative semantic analysis of DRM lists can be found in Brainerd, Yang, Howe, Reyna, and Mills (2008b).] When a DRM list is presented for study, subjects are exposed to the forward associates of the generating word but not to the generating word itself. That word, which is called the critical distractor or critical lure, is used to measure false memory on recall or recognition tests. In individual experiments, subjects are exposed to several such lists, and the modal result is remarkably high levels of false memory for critical distractors: For the lists generated by *anger*, *chair*, *doctor*, and *sweet*, for instance, the mean level of false recall is 54%, and the mean level of false recognition is 76% in young adults (see Roediger, McDermott, & Gallo, 2001). It is difficult to over emphasize just how powerful this illusion is: *Immediately after* exposure to the list *table*, *sit*, ..., *bench*, 54% of college students, on average, recall that *chair* was on the list, and remarkably, they usually recall realistic details of its “presentation”—such as the gender of the voice in which *chair* was “pronounced” or

the font in which *chair* was “printed” (Brainerd, Wright, Reyna, & Mojardin, 2001; Payne et al., 1996).

Turning to back-end manipulations, this is the realm of relative accessibility of verbatim and gist traces, which can be influenced by the retrieval cues that are provided on memory tests and by differential forgetting rates. The general principle, naturally, is that false memory should increase as a function of variables that, at the time memory is tested, enhance gist retrieval relative to verbatim retrieval. Increasing the delay between subjects’ exposure to target events and the administration of memory tests is a frequently studied example, owing to many classic experiments showing that the ability to access memories of the surface details of events degrades more rapidly than the ability to access memories of their semantic content (e.g., Gernsbacher, 1985; Kintsch, Welsch, Schmalhofer, & Zimny, 1990). Thus, if memory tests are delayed for a few hours or days following events, an interval during which verbatim decline is rapid but memory for meaning content is relatively stable, semantic false memory ought to increase because gist retrieval comes to predominate—a prediction that has been confirmed for many types of meaningful events (for illustrations, see Brainerd, Reyna, & Estrada, 2006; Gallo, 2006; Loftus, Miller, & Burns, 1978; Payne, Elie, Blackwell, & Neuschatz, 1996; Reyna & Kiernan, 1994, 1995; Seamon et al., 2002). Actually, even shorter delays can produce dramatic differences in false memory when events are sufficiently complex—sentences in narratives, for instance—that verbatim memories are difficult to retain. Here, Brainerd and Mojardin’s (1998) developmental reversal study is a case in point. In an earlier study, Reyna and Kiernan (1994) detected the standard age decline pattern in children’s false memory for sentences (*The tea is cooler than the coffee*) that preserved the meaning of sentences that children had actually heard (*The coffee is hotter than the tea. The tea is hotter than the cocoa. The cocoa is sweet.*). In their procedure, the accessibility of verbatim traces of sentences was maximized by administering test items immediately following the presentation of each three-sentence narrative. Brainerd and Mojardin used the same procedure, except that they increased the delay between sentence presentation and memory tests to roughly two minutes. The result was an age increase in false memory rather than a decrease.

With respect to retrieval cues on memory tests, the familiar principle of encoding specificity (Tulving & Thomson, 1971) can be exploited to identify cues that ought to shift retrieval in either a verbatim direction or a gist direction. According to encoding specificity, presenting retrieval cues that reinstate the surface features of events (e.g., the gender, pitch, and accent of the voice in which word lists or narratives were spoken), increases the match between those cues and the verbatim traces that contain representations of such features.

Consequently, retrieval cues that closely reproduce the surface features of events should reduce false memory, which they do (Arndt & Reder, 2003; Brainerd, Reyna, & Kneer, 1995; Brainerd, Wright, Reyna, & Payne, 2002; Reyna & Kiernan, 1994, 1995), by enhancing verbatim retrieval. On the other hand, encoding specificity predicts that retrieval cues that reinstate salient aspects of the meaning of events will have the opposite effect, by increasing the match between those cues and the content of gist traces. This prediction has also been confirmed in various experiments (e.g., Lampinen, Copeland, & Neuschatz, 2001), with simple instructions that encourage subjects to relax the tendency to emphasize vivid recollection of surface details when responding to memory tests being quite effective (e.g., Koriat & Goldsmith, 1994, 1996; Payne et al., 1996).

Now that some theoretical principles are in hand, along with a few simple manipulations that implement them, we return to the question of developmental reversals in false memory. The theoretical principles say that other things being equal, variability in the rates of verbatim and gist retrieval determine variability in levels of false memory. With connected-meaning tasks, the extensive literature on the development of semantic processing in memory (for

reviews, see Schneider & Bjorklund, 1998; Schneider & Pressley, 1997) supplies compelling reasons for anticipating that such tasks will reveal parallel increases in false memory. At a minimum, though, the theoretical principles tell us that the standard age decline pattern cannot be the whole story unless some very special conditions are met that we already know are violated by normal development. Specifically, the pattern would be the whole story only if the suppressive side of opponent processes (the verbatim component) waxes between early childhood and young adulthood, while the supportive side (the gist component) either wanes or remains invariant. Although the first condition is true, the second is violated by most of what we know about semantic development and its influence on memory—specifically, that the tendency to store and retrieve the semantic features of events and to form semantic connections among events improve massively between early childhood and young adulthood.

Classic illustrations of that trend can be found in the developmental literature on semantic organization in recall, which shows large increases in spontaneous memory for meaning content and meaning relations when children are exposed to lists of familiar words or pictures of familiar objects (for reviews, see Bjorklund, 1987; Reyna, 1996). Increases in memory for an especially important form of meaning content, taxonomic relations, has been thoroughly documented. When adults study picture or word lists on which the individual items belong to some standard taxonomic categories (e.g., clothing, furniture, fruit, vehicles), their recall exhibits three characteristic effects that suggest that they are storing strong gist traces of taxonomic relations and using them on recall tests to output items (Brown, Flores, Goodman, & Conover, 1991). (a) Total recall is better with such lists than with unrelated lists composed of items of comparable difficulty. (b) The output of items during free recall is clustered by category, as though subjects say to themselves: “There were clothing, furniture, fruit, and vehicles on the list. Let’s see how many items of clothing I can get, then I’ll try furniture, then fruit, and then vehicles.” (c) Such semantic clustering is good for recall because total output increases as the amount of clustering increases. The remarkable thing about memory development is that during the preschool and early childhood years, children exhibit very little evidence of these elementary semantic effects, even though lists are composed of items whose meanings they know and that they can identify when asked (Bjorklund, 2004; Bjorklund & Hock, 1982; Bjorklund & Jacobs, 1985; Bjorklund & Muir, 1988; Chi & Ceci, 1987). The overall pattern is that unless young children are explicitly prompted to store and retrieve taxonomic content, these three effects emerge very slowly during childhood, and they continue to strengthen through adolescence into young adulthood.

In short, based on available data, the known picture of developmental change in verbatim and gist memory is that availability/accessibility of both improves with age (Bouwmeester, Vermunt, & Sijsma, 2007; Brainerd & Reyna, 2004; Reyna, Holliday, & Marche, 2002). The key implication, which is surprising in light of the extensive published evidence of age declines in false memory, is that whether false memory declines, increases, or remains fixed across age levels *will be highly task dependent* but in theoretically predictable ways. Remember in this connection that we have already seen that the mix of experimental conditions on different tasks affects the relative availability/accessibility of verbatim and gist traces. This means that some tasks will be more sensitive to subject variation in verbatim memory whereas others will be more sensitive to subject variation in gist memory. Ontogenesis is a prime source of both types of variation. If the developmental baseline is one of age improvements in availability/accessibility of both types of traces, the standard pattern of age declines in false memory is favored in tasks that maximize verbatim sensitivity while maximizing gist sensitivity (Brainerd & Reyna, 1998; Ceci & Bruck, 1998), but the opposite pattern is favored in tasks that minimize verbatim sensitivity while maximizing gist sensitivity (Brainerd, Reyna, & Forrest, 2002). In other words, when

specific tasks are highly sensitive to variations in verbatim memory but not to variations in gist memory, declines in false memory will predominate because performance is chiefly controlled by variations in verbatim memory, but when verbatim-gist sensitivity is reversed, increases in false memory will predominate because performance is chiefly controlled by variations in gist memory.

Connected-meaning paradigms, the DRM illusion in particular, are rather unambiguous examples of high-gist/low-verbatim tasks because, as we saw, they strengthen gist memory relative to verbatim memory, and they make it difficult to use verbatim retrieval to suppress false memories. Therefore, Brainerd et al. (2002) proposed that these paradigms were obvious places to initiate the search for robust, replicable developmental reversals in false memory that, once identified, could be used to test theoretical hypotheses about processes that control age trends in false memory. That proposal was grounded in two considerations that we have discussed—namely, that (a) children's known limitations in storing and retrieving even simple semantic content imply that their performance on connected-meaning tasks will not be dominated by the strong gist memories that dominate adults' performance and (b) these tasks minimize the effectiveness of verbatim memory in suppressing errors (so that age improvements in that sphere will not override the influence of improvements in gist memory). These considerations also suggest that it should be possible to tie developmental reversals in false memory directly to two general classes of experimental manipulations: *sufficiency* manipulations, which are variables that selectively increase false memory in younger subjects (thereby shrinking the size of developmental reversals) by compensating for their limitations in gist memory and *necessity* manipulations, which are variables that selectively decrease false memory in older subjects (thereby also shrinking the size of developmental reversals) by interfering with their superior gist memory abilities. We shall have more to say on both heads in the third section of this paper, after we sketch accumulated developmental reversal findings in the next section.

Finally, the theoretical principles discussed above are ideas about memory, obviously. In early developmental research on one of the two basic varieties of false memory, the suggestion-induced form, non-memorial concepts—specifically, ideas about age differences in compliance and in motivation to comply with external direction—were also considered as possible sources of children's false memories. In this paper, we do not consider such processes as potential sources of developmental reversal effects, for two reasons. First, the traditional view of compliance and motivation to comply is that they are elevated in younger children, which predicts age *decreases* in false memory (Ceci & Bruck, 1993). Second, a standard finding from much research on false memory is that motivational and compliance factors account for small amounts of variance, relative to memory factors (for a review, see Brainerd & Reyna, 2005).

Core Lines of Evidence

We turn now to the empirical meat, a synopsis of the results of studies that have looked for child-to-adult increases in false memory, using connected-meaning paradigms. Throughout, we will refer to information in Table 1, which contains all the recent examples that we could locate of developmental reversal articles that met two conditions—namely, that the articles were written in English and that they were published in peer-reviewed journals. Even with those constraints, it can be seen that the accumulated evidence of age increases in false memory is quite extensive, consisting of 51 journal articles by investigators from several countries (Australia, Canada, Holland, Portugal, United Kingdom, United States) who tested memory in multiple languages (Dutch, English, French, Portuguese). As can also be seen, only two articles report research failed to detect age increases (null age effects were reported in both), while age increases were reported in all of the 49 remaining articles. Further, some

of the latter articles reported more than one experiment, so that the number of individual experiments in which developmental reversals have been detected is more than 49.

Our summary of the developmental reversal data base proceeds in two steps. First, in the present section, we will concentrate on age trends; on specific findings of increases in false memory between early childhood and young adult. This part of the summary will be organized by experimental paradigm. Referring to Table 1, although developmental reversals have been produced with various paradigms, some clear organizing themes pop out, the most obvious being frequency of use of certain paradigms. Here, it is the DRM illusion versus all other procedures: Scanning down the third column, which describes the procedure that was implemented in each article, it is seen that the DRM illusion was used in 35 articles while other procedures were used in 23 articles. (The sum exceeds 49 because multiple procedures were used in some articles.) The next most common procedure, which was used in 8 articles, consists of presenting sets of categorized materials (words or pictures) to children and measuring false memory for unrepresented exemplars of presented categories (e.g., testing false memory for *piano* and *drums* after studying a word list on which *guitar*, *violin*, *trumpet*, *clarinet*, *trombone*, *oboe*, *saxophone*, and *tuba* appeared). The remaining 15 procedures consist of a miscellany of methods, such as lists of phonologically or emotionally related words, sentences and narratives, reasoning problems, and videos of crimes. To summarize developmental reversal findings, we follow this frequency pattern—beginning with the DRM illusion and devoting the most space to it, then continuing with experiments that used categorized materials, and ending with results from the miscellaneous remaining tasks.

The second step in our review, which we postpone until the section after this, focuses on testing theoretical hypotheses—more explicitly, on results that have been produced by manipulations that embody specific processes, availability/accessibility of verbatim and gist memories in particular, that are thought to foment or to suppress children's false memories. Here, we use the necessity-sufficiency mode of organization; that is, we consider variables that, theoretically, ought to *reduce age increases in false memory* either by interfering with older subjects' superior gist memory abilities (necessity manipulations) or by providing cognitive prosthetics for younger subjects' weaker gist memory abilities (sufficiency manipulations). It will be seen that both types of manipulations have the predicted effects, but we are getting ahead of our story.

Developmental Reversals in the DRM Illusion

By itself, the DRM illusion provides an impressive existence proof that semantic false memories can increase dramatically between early childhood and young adulthood. Remember that in this procedure, subjects are exposed to short lists of words (or, sometimes, pictures; Israel & Schacter, 1997), all of which are forward associates of a missing word that serves as a false memory item on recognition or recall tests. Remember, too, this procedure produces quite high levels of false memory in adults, so that the question is whether it produces even higher levels in children, as would have been expected several years ago, or whether it produces lower levels. It is convenient to divide developmental studies of the DRM illusion into two epochs, 2002–2005 and 2006–present. Because developmental reversal findings are counterintuitive, it is natural to be skeptical for a time about initial findings of that sort; 2002–2005 is that period. It begins with a report of three experiments in which the DRM illusion was found to increase between age five and young adulthood and net memory accuracy was found to decline, and continues with four further articles in which age increases in the illusion were repeatedly replicated and age increases in false memory for categorized materials were also detected. Once researchers had convinced themselves that the developmental reversal pattern was real, the next question was to investigate its

limitations and to ask, in particular, whether it was due to uninteresting methodological variables (such as linguistic differences between children and adults). That period is 2006 to the present, in which a number of potential limiting factors have been investigated, and along the way, 30 more articles have been published in which false memories were found to increase with age.

2002–2005

Brainerd et al. (2002) reported three experiments in which the recall and recognition versions of the DRM illusion were studied in 5-, 7-, and 11-year-olds and in young adults. In the first experiment, 5-year-olds listened to 10 such lists, attempting to recall each immediately after it was presented. The rate of false recall of critical distractors such as *anger*, *chair*, *doctor*, *sweet*, and so on was only 6%. This experiment was replicated point-for-point in a second one, except that (a) 7-year-olds as well as 5-year-olds participated, (b) the number of DRM lists was increased to 16, and (c) both strong and weak DRM lists were administered. Concerning c, in adults, strong lists are ones that produce very high levels of false recall and false recognition of their critical distractors, such as the ones that we have been using as examples, whereas weak lists are ones that produce much lower error rates, such as the lists for the critical distractors *cottage*, *king*, *long*, and *trouble*. However, the children in the experiment, like those in the first, simply displayed near-floor levels of false recall ($M = 7\%$), and false recall did not differ between the two age levels, though *true* recall increased from 31% in 5-year-olds to 38% in 7-year-olds. An especially instructive result of both experiments is that the nature of the items that the children falsely recalled was radically different than in adults. Adults' errors are of two predominant sorts. Most errors are critical distractors (e.g., *chair*), while the remainder ones are overwhelmingly words that, like critical distractors, share meaning with list words (e.g., *bookcase*, *TV*); that is, false recall is almost wholly *semantic*. Not so with young children. Brainerd et al. found that only about half of their errors were semantically related to list words. In their third experiment, 5-year-olds, 11-year-olds, and young adults studied and recalled the same 16 lists as in the second experiment, and they also responded to a recognition test after all the lists had been recalled. Across this age range, false recall of critical distractors quintupled for strong lists and doubled for weak lists, while false recognition increased by 22% for strong lists.

These basic patterns were confirmed in experiments reported in three articles that appeared two years later. In one, Dewhurst and Robinson (2004) used a design like that of the first experiment of Brainerd et al. (2002) in which 5-, 8-, and 11-year-olds listened to 5 short DRM lists (8 words), attempting to recall each immediately after it was presented. False recall of critical distractors roughly doubled with age. Like Brainerd et al., Dewhurst and Robinson found that children's false recall was not dominated by semantic intrusions, as adult's is: The most common intrusions in 5-year-olds were words that rhymed with list words, and it was not until age 11 that semantic intrusions predominated. In the second article, Howe, Cicchetti, Toth, and Cerrito (2004) measured false recall and false recognition of critical distractors in 6-, 9-, and 12-year-olds. A novel feature of their design was that the subject sample included maltreated as well as non-maltreated children, the idea being to test a hypothesis that was common in the clinical literature at the time (see Brainerd & Reyna, 2005)—namely, that maltreatment makes children especially prone to false memories. Consistent with prior developmental studies of the DRM illusion, Howe et al. found that false recall more than doubled with age, from 11% to 24%, while false recognition increased more modestly (from 65% to 74%). The data did not confirm the hypothesis that maltreated children are at higher risk of false memories—at least, not of the type that are measured by the DRM illusion. In the third experiment, Brainerd, Holliday, and Reyna (2004) measured false recognition of critical distractors in 7-, 11-, and 14-year-olds, finding that such errors increased with age from 41% to 68%. We mentioned that a striking

fact about the DRM illusion is that false memories of critical distractors provoke illusory vivid mental reinstatement of their prior “presentation,” which is called *phantom recollection*. Brainerd et al. found that it was these especially compelling false memories that accounted for most of the developmental reversal effect; that is, false memory not only increased with age but it was the most powerful variant that increased.

Summing up the initial group of developmental reversal studies, four general results emerged that have been replicated in many subsequent experiments, so we list them here by way of an interim report. First and most fundamentally, semantic false memory, in the form of the DRM illusion, was found to increase, not decline, between age 5 and young adulthood, regardless of whether it was measured via intrusions on recall tests or false alarms on recognition tests. Second, much like findings in classical studies of the development of semantic organization in recall, false memory varied throughout this age range—rather than, as is often the case in memory development, only varying between early childhood and early adolescence. In short, the ontogenetic picture for the DRM illusion is one of slow maturation. Third, the picture is somewhat different for recall than for recognition. Although both increase with age, false recall seems to be virtually at-floor in early childhood and not to begin moving upward until after age 7, whereas false recognition is above-floor from the start. Fourth, examination of false recall, in particular, reveals just how different memory distortion processes are in children versus adults. In adults, virtually all recall errors are semantic, in that they share meaning with words presented on study lists (Payne et al. 1996). In young children, however, most errors are nonsemantic, consisting of items that appeared on previously studied lists (Brainerd et al., 2002) or items that sound like items on just-studied lists (Dewhurst & Robinson, 2004).

2006-Present

It can be seen in Table 1 that since 2005, 34 articles have appeared in which the DRM illusion produced the developmental reversal pattern. As in earlier studies, developmental reversals have been detected with both recognition and recall. Glancing at the last two columns of Table 1, which report methodological details and key results, it can be seen that these more recent studies have broadened the empirical base in key respects. For instance, developmental reversals have been confirmed in languages other than English—specifically, Dutch (Bouwmeester & Verkoeijen, 2010), French (e.g., Howe, Gagnon, & Thouas, 2007), and Portuguese (e.g., Carneiro, Albuquerque, Fernandez, & Esteves, 2007). In the remainder of this section, we discuss two other noteworthy ways in which the developmental reversal pattern has been broadened.

Perhaps the most important one is that a major methodological explanation of why the DRM illusion increases with age has been ruled out: word comprehension. Brainerd et al. (2002) noted that although the critical distractors that are used to generate DRM lists are familiar words, some of the words on the lists themselves (e.g., *stethoscope* on the *doctor* list) might be too unfamiliar to children for them to be able to make the semantic connections that foment false memories, which means that developmental reversals might be rather uninteresting consequences of age differences in word comprehension. Brainerd et al. rejected this explanation on two grounds. First, they reported that very few words on DRM lists could be considered even moderately unfamiliar. They examined the words’ familiarity scores on the Toglia and Battig (1979) norms—where familiarity is rated on a 1 (lowest) to 7 (highest) scale—and found that the average level of familiarity, 6.23, was very high. Second, Brainerd et al. separated the lists that were administered to their subjects into ones that contained only very familiar words versus ones that contained one or more moderately unfamiliar words and then compared development trends for the two groups of lists. Developmental reversals did not differ in magnitude for the two groups.

However, other investigators—notably, Anastasi, Lewis, and Quinlan (2008), Carneiro et al. (2007), and Metzger, Warren, Price, Reed, Shelton, and Williams (2008)—argued that developmental reversals might still be due to differences in word knowledge of a subtler variety. These researchers pointed out that the DRM illusion relies on word lists that are generated from *adult* word association norms. They hypothesized that child norms—the words that come to *children*'s minds when they are cued with words such as anger, chair, doctor, sweet, and so forth—could be quite different than the words that come to adults' minds. If so, that could explain the lower levels of false recall and false recognition in children, and importantly, age increases in false memory might disappear and the standard age decline trend might be restored if DRM lists were constructed from child association norms. This hypothesis has now been evaluated in multiple ways and disconfirmed in each instance. The developmental reversal pattern is therefore not due to administering adult-normed lists to children.

The first experiment to demonstrate this, though it was not the first to be published, was by Metzger et al. (2008). First, these authors presented some of the critical distractors for adult DRM lists to children, asking them to state words that came to mind. The resulting free associates were then used to construct a set of “child appropriate” DRM lists. Second, a developmental DRM study was conducted in which those lists were administered to 7- and 10-year-olds and adults. Adult-normed lists were also administered to the same age levels. With adult-normed lists, it was again found that false recall increased with age (from 2% to 16%) and so did false recognition (from 23% to 72%). However, contrary to the hypothesis that these age increases are due to developmental differences in word comprehension, the same pattern was obtained with child-normed lists, with false recall increasing from 1% to 13% and false recognition increasing 28% to 46%.

An even more elaborate test of the same hypothesis was conducted by Carneiro et al. (2007). These authors, like Metzger et al. (2008), created child-normed DRM lists (in Portuguese) by first generating child association norms for 16 critical distractors, but Carneiro et al. (2007) generated those norms separately for four different age levels—4-, 7-, 12-, and 24-year-olds. This allowed them to construct four *different* age-appropriate versions of each of the 16 lists, one for each age level. Then, in a developmental study, subjects at each age level were exposed to the specific lists that had been generated by subjects of their age. These more elaborate controls for word comprehension produced the same key findings as Metzger et al. reported—explicitly, that contrary to the word comprehension hypothesis, there was no evidence that false recall and false recognition of critical distractors decreased with age and, instead, the developmental reversal pattern was still present. That pattern was less robust than in prior studies in that age increases in false recall and false recognition were only reliable between the ages of 4 and 7. Because age increases were less robust, some might say that although developmental reversals cannot be entirely due to developmental differences in word comprehension, such differences amplify developmental reversals. However, that conclusion is not supported by Carneiro et al.'s data. Their less robust age trends were probably due to a restricted range problem among their older subjects. With English DRM lists, as mentioned, adults show very high levels of false recall and false recognition, but with Carneiro et al.'s Portuguese lists, adults showed low levels of false recall and false recognition. Thus, the lists that they used did not provide the same statistical power to detect increases at later age levels that English lists provide.

Anastasi et al. (2008) conducted a study that was similar to Metzger et al.'s (2008) in that they constructed a single set of child-appropriate DRM lists by generating a single set of child association norms using 12 critical distractors from adult DRM lists. Six of the child-normed lists and six adult-normed lists were then administered to 5-year-olds, 8-year-olds, and a group of young adults, with subjects responding to a free recall tests immediately

following each list and then responding to a recognition test for all 12 lists at the end of the experiment. Remember, here, that the word comprehension hypothesis predicts a List Type X Age Trend interaction such that developmental reversals will be observed for the adult-normed lists but the standard age decline pattern will be observed for child-normed lists. However, Anastasi et al. found developmental reversals for both types of lists. The increase in false recall was from 23% to 33% with adult-normed lists and 12% to 33% with child-normed lists, while the increase in false recognition was from 45% to 67% with adult-normed lists and from 36% to 53% with child-normed lists.

The second example of how recent studies have broadened the developmental reversal pattern is concerned with alternative definitions of “development” that are based on measured cognitive ability rather than chronological age. Referring again to Table 1, relevant data are reported in the articles by Brainerd, Forrest, Karibian, and Reyna (2006) and Weekes, Hamilton, Oakhill, and Holliday (2007). It will be remembered that a theoretical cornerstone of developmental reversal predictions about connected-meaning tasks is that the ability to extract meaning from items and to connect it across items that share meaning evolves throughout the child-to-young adult age range. Thus, developmental reversals should actually be tied to variations in this ability, rather than to age per se (see also, Ceci, Papierno, & Kulkofsky, 2007). There are various ways to investigate that possibility.

One approach is to study the DRM illusion in samples of learning-disabled children versus samples of nondisabled children who have been equated on variables other than their learning-ability classifications. This is an instructive comparison because children who have been legally classified as learning-disabled must show below-average performance in a school subject (most often, language or reading), but they cannot be below-average in psychometric intelligence. The performance of learning-disabled versus nondisabled children on memory tests has been investigated for many years, and it is well known that learning-disabled children tend to perform poorly when such tests tap the ability to extract and remember meaning content (as in the earlier illustrations of memory for categorized lists). This suggests that semantic processing limitations that are supposedly central to age increases in the DRM illusion are common in learning-disabled children. This, in turn, suggests that the DRM illusion should be weaker in learning-disabled children than nondisabled children at given age levels. That prediction was confirmed in some research reported by Brainerd et al. (2006).

These authors administered a series of DRM lists to 7- and 11-year old children. Half the subjects at each age level were children who had been legally classified as language- or reading-disabled and were receiving school services for their disabilities, while the others were nondisabled children who were not receiving any special services and were matched on IQ. All of the children were exposed to a total of 16 DRM lists and were asked to recall each list immediately after its presentation. In the sample as a whole, false recall of critical distractors increased from 11% at age 7 to 24% at age 11. As predicted, false recall was lower in learning-disabled children than in nondisabled children among 7-year-olds (6% versus 16%) and among 11-year-olds (19% versus 30%).

Another, more precise, approach to tying the DRM illusion to developmental changes in semantic processing ability was taken by Weekes et al. (2007). Following the above line of reasoning, these authors predicted that the DRM illusion ought to be weaker among children who show below-average performance on standardized tests that specifically measure their ability to understand the semantic relations that exist between common words. For instance, that ability is the focus of tests of reading readiness and reading comprehension. Those are the types of tests that Weekes et al. administered, using them to construct a sample of

children, half of whom displayed reduced performance. Explicitly, Weekes et al. administered a series of DRM lists to two groups of 9- to 11-year-old children, with each list being recalled immediately following presentation. The children in one group performed a year or more below national age norms for these tests, but they performed at the norms for their age on tests for *nonsemantic* reading abilities (e.g., phonology) and on intelligence tests. The other group of children performed at the norms for their age on all tests. Weekes et al. found group differences that were precisely localized within false memory: Children with low versus normal semantic processing ability displayed the same levels of true recall, but as predicted, false recall was greatly reduced in the low group as compared to the normal group. Another instructive feature of Weekes et al.'s research is that this group difference was narrowly localized within *semantic* false memory. In addition to standard DRM lists, Weekes et al. also administered nonsemantic analogue lists that were developed by Sommers and Lewis (1999). In the analogue task, the lists are composed of words (*hat, rat, sat, that, cab, cot, caught*, and so on) that are all phonologically related to a familiar critical distractor (*cat* in this instance). With such lists, there were *no differences* in false recall between the two groups of children, so that the group difference that was observed with standard DRM lists was closely tied to group differences in semantic processing.

Afterword

The developmental database on the DRM illusion supplies an existence proof that semantic false memory can increase substantially during childhood and adolescence, contrary to the law's traditional assumptions about the developmental course of memory distortion. More than three dozen articles have been published, using subject samples from multiple countries and testing the DRM illusion in languages other than English. Two well-replicated findings are: (a) False recognition and false recall of unpresented meaning-sharing words increase between childhood and adolescence and between adolescence and young adulthood, and (b) false recall is at-floor in early childhood while false recognition is above-floor. Another result of forensic moment is that in some articles (e.g., Brainerd et al., 2002; Metzger et al., 2008), age increases in false memory have outstripped corresponding increases in true memory—so that net accuracy (defined as the probability of remembering a list word divided by the probability of remembering a list word plus the probability of remembering the critical distractor) actually *declines with age*.

Although experimental support for developmental reversals is massive, there is a wholly predictable criticism that is often voiced in forensic circles. That criticism is that data on the DRM illusion are irrelevant to legal cases because they lack ecological validity. After all, the criticism goes, legal cases deal with memory for complex real-life events, but the DRM illusion is a word-list task. This is a variant of a hoary criticism of word-list tasks that is traditionally raised by everyday memory researchers (see Banaji & Crowder, 1989). Although the objection seems utterly convincing to many, it is flawed in two fundamental ways. First, note that it is in the nature of proof-by-assertion. The criticism is offered as a self-evident proposition rather than as a conclusion that grew out of data. Rather than being swept along by rhetoric, we need to remind ourselves that the history of science is littered with experimental disconfirmations of seemingly self-evident ideas that were accepted for hundreds of years—Galileo's demonstration that heavy and light objects actually fall at the same rate being a familiar illustration.

Second, when we consider whether experimental evidence can be located to support the criticism, we find that the data all are on the other side of street. With respect to the DRM illusion, in particular, Gallo (2011) showed in a recent literature review that individual differences in the illusion are reliable predictors of distortions in everyday autobiographical memory and more exotic autobiographical distortions, including recollections of living past lives and being abducted by aliens and adult recovered memories of previously

unremembered childhood abuse. Further, in the social psychology literature, the DRM illusion has been found to predict false memories of complex social situations (e.g., Garcia-Marques, Ferreira, Nunes, Garrido, & Garcia-Marques, 2010). In short, the above claim, seductive though it may be, is simply wrong empirically; there are ample data connecting the DRM illusion to distortion of memory for complex, real-world experiences.

It is important to add that the overriding objection to word list data, which the ecological validity criticism of the DRM illusion devolved from, is also wrong empirically. At a general level, Banaji and Crowder (1989) and many others before us have commented that just about every basic law of human memory—laws that are routinely applied in courtrooms, clinics, and classrooms—was discovered with word-list tasks (e.g., the forgetting function, reminiscence, testing effects, encoding specificity, massed versus distributed practice, proactive and retroactive interference, short-term memory capacity, serial position curves). If such tasks are irrelevant to memory for complex real-life events, why have they generated general laws of human memory that are instantiated repeatedly in real-life contexts (Reyna et al., 2007; Rohrer & Pashler, 2010)?² At a specific level, performance on word-list tasks has been repeatedly tied to a variety of clinical conditions in which memory distortion is an important concomitant or even the prime presenting symptom. The two most common forms of neurocognitive impairment in older adults, Alzheimer's dementia and its precursor condition, mild cognitive impairment, are cases in point. Although a variety of neuropsychological tests, medical tests, and everyday functioning information are used to diagnose these conditions, the best single predictor of such diagnoses is simple recall of word lists (see Petersen, 2004). Likewise, in young adults, simple recall of word lists differentiates individuals with schizo-typic symptoms, who are at risk of developing schizophrenia, from individuals who are not at risk (Brainerd, Reyna, & Howe, 2009). As a third example, memory distortion is a characteristic of various emotional disturbances, such as post-traumatic stress disorder, so that it is essential to secure objective measurements of emotional reactions to stimuli in individuals with such conditions. Word lists provide a convenient, clinically useful technique for obtaining those measurements because it has been found that they produce emotional reactions that parallel those that are produced by real-life events and autobiographical memories (Rubin & Talacrico, 2009). Empirical demonstrations of this ilk could be multiplied indefinitely, and we will consider another one later, when we take up developmental reversals in emotional false memory.

Developmental Reversals in False Memory for Categorized Materials

In one of the early DRM articles, Brainerd et al. (2004) also reported an experiment on age trends in false memory for categorized materials. The object was to generalize DRM developmental reversals by seeking the same pattern with a different connected-meaning paradigm that had the properties that, according to FTT, foment developmental reversals. In this alternative procedure, meaning connections are established via taxonomic relations, rather than norms of word association, and false memory is measured via intrusions and false alarms for unrepresented events that are taxonomically related to presented events. The most common version of this procedure is a word-list task in which subjects are exposed to a short list consisting of exemplars of a single taxonomic category (e.g., *hand, nose, neck, mouth, stomach, knee, heart, brain, chest, shoulder, toe, leg, thigh, ankle, face*) or to a longer list composed of several exemplars of each of a small number of categories (say, 8 exemplars apiece from the animal, clothing, fruit, and furniture categories). With either type

²Some commentators have been concerned that modulation of the laws of memory challenges the validity of such laws. However, the laws of physics are modulated by a host of factors when they are applied to the real world (e.g., friction with the road affects the inertia of a car, yet the laws of physics still apply to cars and are applied in accident analyses).

of list, there will be several familiar exemplars that were omitted that can serve as false memory items; for instance, *ear*, *eye*, *finger*, and *foot* for the body part list.

Brainerd et al. (2004) exposed 5- and 11-year-old children to three lists of category exemplars (e.g., animal names, clothing names, and color names) followed by a recognition test on which the presented exemplars of each category, unpresented exemplars of each category, and exemplars of unstudied categories were all tested. The middle group was the false memory items, naturally. In all, there were three cycles of three categorized lists followed by a recognition test, for a total of nine categorized lists. There was a developmental reversal effect: False recognition of unpresented exemplars of studied categories increased by more than 40% between the ages of 5 and 11.

Shortly thereafter, Howe (2006) extended this finding to false *recall* of unpresented category exemplars. In his research, 5-, 7-, and 11-year-olds were exposed to 8 14-item lists, with each being composed of exemplars of a single taxonomic category. From each of these lists, Howe omitted the most frequently mentioned exemplar of that category and designated it as the critical distractor. After listening to each list, children participated in a 30 sec distractor activity, to empty short-term memory, and were then told to recall as many items from the list as possible. Howe found that false recall of critical distractors doubled over this age range, increasing from 16% to 32%.

Sloutsky and Fisher (2004; Fisher & Sloutsky, 2005) implemented a procedure that is different from anything we have considered so far, but they obtained the same developmental reversal pattern. As with categorized word lists, the target materials were all exemplars of a single taxonomic category, but those targets were so physically similar as to make it exceptionally difficult for subjects of any age to use verbatim retrieval to reject unpresented exemplars of the category. More specifically, the materials consisted of several color photographs of each of three types of animals (e.g., several pictures of bears, several pictures of birds, and several pictures of cats). When the pictures were presented for viewing, those from one of the categories (e.g., cats) were tagged with a meaningless verbal cue (“has beta cells”) that differentiated exemplars of that category from exemplars of the other two categories. Later, a recognition test was administered that contained presented pictures from each category, unpresented pictures from the tagged category, and unpresented pictures of animals from unpresented categories (e.g., squirrels), with the second group of pictures being the false memory items. In an initial developmental study (Sloutsky & Fisher, 2004), this task was administered to a sample of 5-year-olds and a sample of young adults, and false memory was found to increase dramatically with age: False alarms to unpresented pictures from studied categories increased from 41% to 76% while hits to presented pictures only rose from 72% to 83%—so that like some other DRM studies, net memory accuracy on this task declined between early childhood and young adulthood.

These patterns were replicated in a more extensive developmental study that used the same task but that included subjects from other age levels—explicitly, 7- and 11-year-olds—as well as 5-year-olds and young adults. Across these four age levels, false alarms to unpresented pictures from studied categories increased from 40% to 45% to 59% to 74%. Thus, the developmental reversal trend was most marked during later childhood and adolescence, which is also the age range during which age increase in the semantic organization of memory are most marked (Brainerd & Reyna, 2004, 2005). Hits to presented pictures increased more modestly over this age range, from 70% to 81%, so that net memory accuracy once again *declined between early childhood and young adulthood*. Thus, when developmental findings from this paradigm are added to findings from some earlier DRM studies (e.g., Brainerd et al., 2002; Metzger et al., 2008), it has become clear that not only can developmental reversals be consistently obtained under the theoretically specified

conditions that we sketched earlier, but those reversals can be so robust that they swamp age increases in true memory, leading to declines in net accuracy.

Returning again to Table 1, it can be seen that since the above four articles appeared, four others have been published in which age increases in false memory were detected for unrepresented exemplars of studied categories. Categorized word lists were used in three of the articles (Brainerd & Reyna, 2007; Dewhurst, Howe, Berry, & Knott, 2011; Knott, Howe, Wimmer, Dewhurst, 2011), whereas the Sloutsky-Fisher picture paradigm was used in the remaining one (Wilburn & Feeney, 2008). In one of the word-list experiments (Brainerd and Reyna, 2007), the false alarm rate for unrepresented exemplars of studied categories increased from 25% to 67% between the ages of 6 and 14, while the hit rate for presented targets only increased from 49% to 82%. In the Wilburn and Feeney (2008) article, these authors, like Sloutsky and Fisher (2004; Fisher & Sloutsky, 2005), found that age increases in false memory were greater than increases in true memory—so that once again, net memory accuracy declined between early childhood and young adulthood.

In sum, studies of false memory for categorized materials have generalized the developmental reversal pattern that was initially detected with the DRM illusion in some fundamental ways. The most obvious one is that robust developmental reversals can easily be obtained with very different types of materials, ones that are generated from category norms rather than association norms and that include realistic pictures as well as word lists. Another important generalization concerns net developmental reductions in memory accuracy; that is, reductions in the proportion of reported items that are in fact true. Some of the earlier DRM studies produced this dramatic form of the developmental reversal, and now, it has been found in all of the experiments that implemented the Sloutsky-Fisher picture methodology and in some of the experiments that used categorized word lists. The fact that net memory accuracy can decline with age is of great forensic interest because it suggests that not only are older witnesses more likely to produce certain types of false memories than younger witnesses, but the overall yield of accurate information (the difference between true and false information) can be lower in adult witnesses. It is difficult to overstate the forensic significance of the finding that *under predictable conditions that apply to real-life contexts*, adults are less accurate than children and older children are less accurate than younger children.

Developmental Reversals in False Memory for Complex and Forensic Events

In our discussion of DRM studies, we aired the familiar criticism that the results of such studies are irrelevant to the reliability of testimony because word-lists lack ecological validity when it comes to the complex, real-world experiences that figure in testimony. We noted that although that criticism strikes many as incisive, it is misguided, for two reasons. First, it is a rhetorical flourish, not an empirical generalization growing out of controlled experimentation that compared false memory for more complex events to false memory word lists and demonstrated that they do, indeed, produce qualitatively different patterns. Second, this criticism has been disconfirmed by data showing (a) that the DRM task predicts false memory for various complex, real-life events, including ones that are characteristic of clinical conditions and (b) that other types of word-list tasks do likewise.

Although data are the final arbiter of truth in science, the history of psychology teaches us that hypotheses with strong rhetorical appeal have a habit of surviving mountains of disconfirmatory results. Therefore, in this section, we answer the ecological validity criticism directly by considering developmental reversal studies of false memory for complex events, including events of obvious forensic relevance, using paradigms that have

often been discussed in expert testimony as supporting the traditional assumption that false memory declines with age. The denouement is that when the theoretical conditions that favor developmental reversals are satisfied (i.e., connect-meaning paradigms), it is not just word lists (or picture lists in the case of the Sloutsky-Fisher paradigm) that produce such reversals; so do complex events. In Table 1, eight articles of that type are listed, which can be subdivided into two groups of articles that we review separately below: three in which the measured false memories were spontaneous (Fernandez-Dols, Carrera, Barchard, Gacitua, 2008; Lyons, Ghatti, Cornoldi, 2010; Odegard, Cooper, Lampinen, Reyna, & Brainerd, 2009) and five in which they were generated by post-event misinformation (Ceci, Papierno, & Kulkofsky, 2007; Connolly & Price, 2006; Fazio & Marsh, 2008; Principe, Guiliano, & Root, 2008; Ross et al., 2006). All of these studies were designed in the wake of the early word-list studies that detected developmental reversals and adopted as one of their objectives to determine whether such reversals also occur for complex events. Thus, they provide direct tests of the ecological validity criticism.

Developmental Reversals in Spontaneous False Memories

The first study of this sort was reported by Fernandez-Dols et al. in 2008, with children in the 6–9 age range. In an initial pair of experiments with adults, the authors validated a procedure for studying false memory for a form of information that is often central in legal cases: people's emotions. In this procedure, subjects were shown either videos or realistic slide sequences containing multiple characters, some of whom were adults and some of whom were children. The characters' faces exhibited either happy expressions or fearful expressions, the authors' objective being to create gist memories of happiness and gist memories of fearfulness. On later recognition tests, subjects viewed further pictures of the faces of the same characters. Some of the pictures had been seen before and some had not, but all of the characters were the same as before. Among the previously unseen pictures, the false-memory items were ones in which the characters' faces displayed the same type of emotion as they had displayed before. This procedure yielded an impressive false memory effect in adults. When it was administered to children, there was a clear developmental reversal: The percentage of subjects who falsely recognized new facial expressions that preserved the emotional gist of videos and slide sequences increased from 75% (age 6–7) to 90% (age 8–9) to 96% (adults).

In the remaining two articles, false memories for scripted events from real life were investigated in children of different ages. In the Lyons et al. (2010) experiment, subjects from five age levels (6-, 7-, 9-, and 10-year-olds, plus young adults) were exposed to picture stories of events from four familiar thematic situations (eating at a restaurant, getting up in the morning, grocery shopping, and attending a class at school). The stories were constructed in such a way that certain events were depicted (e.g., a pile of oranges scattered on the floor of the vegetable section of a grocery store) that must have been caused by another event (e.g., a shopper taking an orange from the bottom of the pile rather than from the top), but that causal event was *not depicted* in any of the pictures. Later, subjects received a picture recognition test containing both old and new pictures and were asked identify the old pictures. Among the new pictures, the false-memory items were ones that depicted the previously unseen causal events. The false alarm rate for these unseen causal events increased from near-floor to roughly 25% in 9-year-olds and remained constant thereafter. Also, when subjects were asked to rate their confidence in the accuracy of these false alarms, confidence increased steadily throughout the 6 to young adult age range.

In the final study, by Odegard et al. (2009), the false memories that were measured were for real-life events in which children (5- to 12-year-olds) directly participated. Those events consisted of participation in birthday parties. To establish the conditions for formation of the types of gist memories that produce developmental reversals, children (a) attended a series

of four birthday parties that were spaced over four consecutive days, and (b) each party revolved around a familiar theme (either the birthday of characters from the Sponge Bob Square Pants television program or of characters from the Harry Potter novels/movies). During each party, children participated in some activities with were directly related to the party's theme and in some unrelated activities as well. Ten days after the fourth party, children received a forensic-style investigative interview about party events, using the well-known NICHD protocol (e.g., Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007; Poole & Lamb, 1998). These interviews included tests for the types of events that have produced developmental reversals in studies that we have reviewed so far (i.e., ones that preserved the themes of birthday parties), as well as the types of events that have not been found to produce developmental reversals (i.e., ones that were unrelated to party themes). Between the ages of 5 and 12, false memory for gist-consistent events increased reliably, but false memory for unrelated events did not.

Developmental Reversals in Implanted False Memories

The Loftus (1975) misinformation paradigm occupies a special place in developmental research on the reliability of child witnesses' memories. The forensic relevance of the procedure is undeniable because it is an experimental analogue of the suggestive, leading lines of questioning that are so prominent in criminal investigation and interrogation, lawyerly questioning of witnesses, and even in everyday conversations about crimes. Consequently, beginning in the 1980s, the misinformation procedure has been extensively used in developmental research. Indeed, it is the principal source of data supporting the traditional developmental decline assumption and, hence, the principal from of scientific evidence on which expert testimony favoring that assumption has been based (McAuliff & Kovera, 2007). Ceci, Ross, and Toglia (1987) were the first to report a series of controlled experiments showing that susceptibility to memory suggestion for everyday events decreased between early childhood and young adulthood. Since then, dozens of follow-up studies, conducted by many researchers in multiple countries and languages, have replicated that pattern and isolated key variables that influence it, such as delay between original events and misinformation, delay between misinformation and memory testing, plausibility of misinformation, credibility of misinformation sources, and so on. The early portion of this literature was summarized in an influential review by Ceci and Bruck (1993), and later work has been summarized in other reviews (e.g., Brainerd & Reyna, 2005; Goodman, 2006; Goodman & Schaaf, 1997; Holliday, Reyna, & Hayes, 2002; Quas, Qin, Schaaf, & Goodman, 1997; Reyna, Mills, Estrada, & Brainerd, 2007).

Despite the massive evidence of age declines in false memories that are implanted with the misinformation paradigm, it is natural to wonder whether the developmental reversal pattern could also be produced if this paradigm were to be modified to meet the theoretical conditions that were sketched earlier; that is, modified to be a connected-meaning task in which it is hard for developmental increases in verbatim memory to neutralize parallel increases in gist-based false memory. It turns out that the answer is yes. Connolly and Price (2006) were the first to publish such a demonstration. In order to create a connected-meaning variant of the paradigm, one group of subjects (preschoolers and first graders) participated in a single play session while a second group participated in four play sessions that established connected meaning relations. During the first session in the connect-meaning condition, each child performed eight target activities that were accompanied by a critical detail, such as a white object. When those activities were repeated in each of the next three sessions, they were accompanied by new critical details that preserved the gist of prior details (e.g., brown, blue, and yellow versions of the *same object*). The final play session was followed by a misinformation phase that focused on these critical details and suggested that other details had been experienced that preserved the gist of the details that had actually

been experienced (e.g., a black version of the white/brown/blue/yellow detail). A day after the misinformation phase, a series of free and cued recall tests for the four play sessions were administered. In both the control and connected-meaning conditions, children exhibited reliable false recall of suggested details. However, in the connected-meaning condition, those errors tripled with age on free recall tests and quadrupled with age on cued recall tests. Thus, the standard developmental decline pattern for implanted false memories was reversed by simply modifying the misinformation paradigm to satisfy the theoretical conditions for age increases in false memory.

The next experiment to demonstrate the same thing was published three months later by Ross et al. (2006) and focused on a question of great forensic importance: the suggestibility of witnesses' memories of the faces of criminal suspects. Similar to the reasoning behind Connolly and Price's (2006) research, Ross et al. thought that the tendency to process any salient semantic relations between culprits and innocent bystanders—such as their age, ethnicity, or gender—should increase substantially with development (on analogy to increases in the tendency to process taxonomic relation in word-list tasks). On that basis, they conjectured that suggestion-induced false identifications of innocent bystanders should exhibit developmental reversals when culprits and bystanders shared such semantic relations. To determine if that was so, they presented a video of a theft of money in a cafeteria to 5-, 7-, 9, and 11-year-old children. Among other characters, the video contained a culprit and a bystander of the same age, ethnicity, or gender. Later, the children received an eyewitness identification test, consisting of five photographs. The photographs were of the innocent bystander and four people who were not present in the video. One of the testing conditions involved misinformation to the effect that the thief was definitely present among the photographs. In that condition, the tendency to falsely identify the innocent bystander as the thief increased dramatically with age, from 18% in 5-year-olds to 40% in 7-year-olds to 47% in 9-year-olds to 64% in 11-year-olds.

A year later, Ceci et al. (2007) reported the next misinformation study in which connected meaning produced a developmental reversal effect. The important advance in this research is that the authors actively manipulated age differences in the tendency to store gist memories of meaning content via a semantic distance procedure. In that procedure, the semantic distances between 22 common objects (e.g., a lemon, an orange, a glass of milk, a cow) were separately determined for a group of 4-year-olds and a group of 9-year-olds. That was done by presenting pictures of the objects in groups of three, designating one picture as the focal object, and asking children to judge which of the remaining two pictures was more similar to the focal object. Using this procedure, it was possible to identify groups of objects that were more similar to older children than to younger ones—the idea being that if older and young children were presented with pictures of those objects, older children would be more likely to connect the gist and, therefore, would be more susceptible to misinformation about unrepresented pictures that share that gist. That is what Ceci et al. found, using a standard misinformation experiment in which (a) a picture story was presented to children that contained eight of the objects whose semantic distances had been rated, (b) children were then exposed to a misinformation phase suggesting that certain objects were presented in the story that were not, and finally, (c) children responded to a recognition test consisting of target objects that appeared in the story, distractor objects for which they had received misinformation, and distractor objects for which they had not received misinformation. The key prediction was that susceptibility to misinformation would increase with age when misinformed objects were more closely connected in meaning for older children than for younger ones. This is exactly the pattern that was obtained, with false alarms rates for such objects doubling between the ages of 4 and 9.

The next misinformation experiment to produce developmental reversals was reported by Fazio and Marsh (2008). To begin, Fazio and Marsh, in prior research with adults, had developed a procedure in which subjects were given false suggestions about real-world “facts,” concerning which they presumably possessed conflicting knowledge, with the false “facts” being cleverly embedded in stories about familiar events, such as wedding anniversaries and the Civil War. Afterwards, they responded to cued recall tests about events in the stories. After reading such a suggestive story only twice, 40% of young adults recalled false facts that conflicted with their own knowledge. Based on FTT’s analysis of developmental reversals, Fazio and Marsh reasoned that if susceptibility to suggestion depends on forming strong gist memories by connecting meaning across different events, these false memories ought to increase with age. Note that this provides a very strong test of the underlying theoretical ideas: Common sense would say that older subjects should be *less susceptible* to such misinformation because they are more likely than younger subjects to possess the factual knowledge that is required to resist misinformation (e.g., that Stonewall Jackson was a Confederate general, not a Union general). However, the data were in line with theoretical prediction rather than common sense in that recall of false suggested “facts” increased from 6% in 5-year-olds to 13% in 7-year-olds and, as mentioned, to 40% in young adults.

The last article in this group of is by Principe et al. (2008), and it involved another form of misinformation that is of high forensic interest: rumor mongering. A standard feature of criminal cases, once they become public, is that rumors circulate back and forth among witnesses, other parties to the crimes, their family members, friends and acquaintances. When suspects are identified, rumor mongering about their past behavior can be especially intense. A natural concern, then, one that has received surprisingly limited attention in research, is whether the gossip and innuendo that circulate during rumor mongering can distort witnesses’ memory for the actual details of crimes. For instance, will witnesses remember defendants committing criminal acts that, actually, they only heard about via rumor mongering? Principe et al. (2008) investigated such questions in children under controlled conditions. The children in her experiment were 3- and 6-year-olds. To ensure that the experiment was a connected-meaning task, Principe et al. studied false memories that like Lyons et al.’s (2010) experiment, involved the causes of effects that had been directly experienced by children. In other words, some of the events that children experienced were caused by other events that they had not experienced, but unlike Lyons’ research, suggestive information about the causal events was provided to children. Specifically, a visitor to the children’s classrooms performed a magic show that involved various tricks, two of which consisted of pulling a rabbit out of a hat and producing a baked cake after simply mixing cake ingredients in a pan. The visitor failed at both of those particular tricks. After the magic show, rumor mongering was induced in children who had been assigned to a suggestion condition by exposing them to a series of indirect hints about what might have happened to cause the hat and cake tricks to fail. (The hints were intended to encourage the inferences that the hat trick had failed because the rabbit had gotten loose in the classroom and that the cake trick had failed because one of the children had eaten the cake without the magician’s knowledge.) To confirm that children in the suggestion condition propagated rumors about why the tricks failed to control children who had not received hints, their behavior was observed by a confederate. Finally, a week later, all the children received a memory interview about the events of the magic show, with chief interest attaching to whether they falsely remembered causal events that were only hinted at in the suggestion condition and that were conveyed to control child via rumor mongering. The children in both conditions exhibited such false memories. More important, there was an Age X Condition cross-over in the development in false memory. For children in the suggestion condition, who like Lyons et al.’s children, had to connect meaning and make causal inferences to generate false memories, false memory for unseen causal events

increased substantially between age 3 and age 6. For children in the control condition, who, like children in standard misinformation experiments, passively received suggestions via rumor mongering, false memory for the same events declined with age.

Comment

Beyond the laboratory paradigms that we reviewed earlier, then, the literature on developmental reversals also supplies direct disconfirmations of the ecological validity criticism by demonstrating age increases in false memory for complex events. That work includes experiments that have used forensically relevant tasks whose data have long figured in expert courtroom testimony. However, there is an important type of research that has not yet been undertaken—namely, studies linking age increases in false memory in these tasks to age increases in the word-list paradigms that were covered earlier in this paper. We mentioned that in the adult false memory literature, there is extensive research showing that false memory in tasks such as the DRM illusion is correlated with a variety of complex, real-world false memories—including exotic and clinical forms (Gallo, 2010). As yet, developmental research of that ilk has not been conducted. It should be, for both theoretical and practical reasons. On the theoretical side, there are straightforward predictions that such research would put to the test: If, as theory assumes, the same processes are responsible for developmental reversals in both spheres, age trends in the two types of tasks should correlate. On the practical side, it is far simpler to design and conduct controlled experiments with laboratory paradigms such as the DRM illusion and categorized lists than with the sorts of paradigms that we reviewed in the present section. Consequently, future researchers' tasks would be simplified if comparability could be established between the two domains, .

Developmental Reversals in Emotional False Memory

We now review evidence for developmental reversals in a type of false memory that is of special forensic interest: false memory for emotional events. A shared property of the legal cases that first drew attention to children's false memories is that child witnesses were attempting to remember emotionally charged events. This is true of most criminal cases, of course. Emotional arousal is intrinsic to some of the events that child witnesses attempt to remember, such as specific abusive behaviors that they experienced or that they saw inflicted on others. Other key events that child witnesses attempt to remember, such as the physical appearance of a culprit or the time of day, are not emotionally arousing in themselves, but the accuracy of children's memories for those events may be influenced for good or ill by the fact that they were placed in a state of emotional arousal by other events that were part of a crime. As emotional states are an integral part of memory in legal contexts, how such states affect levels of false memory is of great importance. A current hypothesis in adult research, for which there is a good deal of recent support (Brainerd, Stein, Silveira, Rohenkohl, & Reyna, 2008c; Goodman et al., 2011), is that emotional states, especially negatively valenced ones, increase the rate of false memory for information that is consistent with those emotional states. The key developmental question then becomes whether this distortive effect decreases or increases with age.

There are two contrasting hypotheses about that. One falls out of a leading adult theory of emotional memory, the affect-as-information hypothesis (e.g., Carstensen & Mikels, 2005; Corson & Verrier, 2007; Isen, 1987; Storbeck & Clore, 2005), while the other falls out of FTT's analysis of the growth of gist memory (Brainerd et al., 2010; Rivers, Reyna, & Mills, 2008; Goodman et al., 2011). According to the affect-as-information hypothesis, adults focus their processing attention on the surface features of experience when they are in negative emotional states, as compared to when they are in neutral or positive states. This means that negative states selectively strengthen verbatim memory for actual events, which

should reduce false memory for gist-consistent events, relative to neutral or positive states (Storebeck & Clore, 2005). Because verbatim memory improves between early childhood and young adulthood, this protective effect of negative emotion should increase with age, causing false memory to decline with age.

The prediction that false memory for negatively valenced experiences can increase with age can be generated from FTT, and it is predicated on a distinction between two core dimensions of such experiences, valence and arousal. Although specific emotional states (e.g., anger, fear, joy, sadness) differ in a number of ways, it is well known that all such states vary along a valence dimension (anger is more negative than joy) and an arousal dimension (sadness is less arousing than anger). Rivers et al. (2008) and Brainerd et al. (2008) pointed out that in FTT, valence is a conceptual aspect of experience that is subject to the usual development trends in memory for semantic gist (see also, Goodman et al., 2011). Negative valence is viewed, on analogy to familiar taxonomic relations, as a salient type of meaning connection among events. Arousal, on the other hand, is nonconceptual and is viewed, on analogy to dual-task inference, as a form of noise that interferes more with verbatim memory than with gist memory. Developmentally, then, the valence side of emotion favors age increases in false memory, especially for negative valence, because, like other meaning relations, the ability to connect valence gist among different events increases with age.

In developmental studies, there are ethical reservations about inducing the intense emotional states that are associated with serious crimes. Thus, most research on this topic should implement milder forms of emotion to tie down the key effects of valence and arousal and how they vary with age. Here, researchers have devised two methodologies, one of which has been adopted in multiple developmental studies (see Table 1). In one method, which is common in the adult social psychology literature (e.g., Corson & Verrier, 2007; Storebeck & Clore, 2005), subjects are placed in a positive or negative mood, then they are exposed to *neutral information* (e.g., a prose passage from a history book), and finally, true and false memory for the neutral information are tested. Various procedures can be used to induce moods, such as listening to appropriate musical passages or viewing a series of pictures from the International Affective Picture Norms (IAPS; Lang, Bradley, & Cuthbert, 2005). [The IAPS contains many pictures of high-arousal negative stimuli (poisonous snakes, vicious animals), high-arousal positive stimuli (joyous faces, people water skiing), low-arousal negative stimuli (elderly people, a physician's office), and low-arousal positive stimuli (ballet dancers, pastoral landscapes).] The other method, which has been used in recent developmental research, is to expose subjects to events that are inherently emotional and simply measure true and false memory for those emotional events.

The particular version of the second procedure that has most often been implemented was introduced by Budson, Todman, Chong, Adams, Kensinger, and Krangel (2006), and it is a variant of the DRM illusion. Using the same association norms that generate standard DRM lists, Budson et al. created a series of parallel lists whose critical distractors were negative emotional words, such as anger, cry, lie, and sick. Howe (2007) then administered those lists, along with emotionally neutral DRM lists to samples of 5-, 7-, and 11-year-olds, and he found on recognition tests that false alarm rates were higher for emotion critical distractors than for neutral ones, and he found that emotional false memory, like the neutral variety, increased with age. Emotional false memory also increased with age with recall tests, but false recall was higher for neutral critical distractors than for emotional ones. Howe, Candel, Otgaar, Malone, and Wimmer (2010) then reported two experiments with adults and three with 8- and 11-year-olds, in which false recall and false recognition were measured with Budson et al.'s emotional and neutral lists. The first two experiments replicated Howe's earlier result that which type of list (emotional or neutral) produced more false memory flip-

flopped for recognition versus recall, and the other three experiments showed that false recall and false recognition both increased with age for emotional as well as neutral lists. The three developmental experiments also produced important new findings because they dealt with changes in emotional and neutral false memory over a one-week delay, using recognition tests. Here, a frequently studied prediction of FTT that was discussed earlier is that false memory levels drift upward over delays of this sort because verbatim traces become inaccessible more rapidly than gist traces. Consistent with that notion, Howe et al. found that true memory declined but false memory increased over the delay, for both emotional and neutral lists and for children as well as adults. Moreover, the upward drift in false memory was greater for emotional than for neutral lists, so that emotional false memory was considerably higher than its neutral counterpart after a week. Age increases in false memory for emotional as well as neutral lists continued to be present after a delay, with levels of false recognition of critical distractors being roughly 20% higher in adults than in children.

Different results were obtained in adolescents and middle-aged adults by Goodman et al. (2011). These authors' administered neutral and negative DRM lists, but they also administered positive ones, and both age levels included subjects who were victims of child sexual abuse. Whereas Howe and associates obtained higher levels of false memory for negative lists with recognition tests but higher levels for neutral lists with recall tests, Goodman et al. obtained higher levels of false memory for negative lists with *recall* tests but higher levels for neutral lists with recognition tests. Also, whereas Howe and associates observed age increases in false memory for both negative and neutral lists, Goodman et al. obtained a null age effects.

The research of Howe and associates and Goodman et al. (2011) seems to show that there are robust false memory effects for the valenced version of the DRM illusion, that the valenced version can be more powerful than the standard version under certain conditions, and that the valenced version, like its neutral counterpart, exhibits developmental reversals. However, there are uncertainties about these findings owing to fact that valence and arousal are confounded in the Budson procedure, and hence, it is impossible to know whether the emotional false memory effects that are observed are caused by valence, arousal, or both. Also, when the relative magnitude of emotional versus neutral false memory is found to flip-flop as a result of the type of memory test (Howe, 2007; Howe et al., 2010), it could be because verbatim memory contributes more to one type of test than to the other.

To eliminate the valence-arousal confound, Brainerd et al. (2010) developed a pool of 32 emotional DRM lists, which are shown in Table 2. There are 16 negative lists, 8 higher-arousal versus 8 lower-arousal, and 16 positive lists, 8 higher-arousal versus 8 lower-arousal. The mean valence ratings (on a 9-point scale) of the 15 list words and the critical distractor are the same for higher-arousal versus lower-arousal lists, and conversely, the mean arousal ratings are the same for negative versus positive lists. Thus, by sampling subsets of the lists, it is possible to conduct studies of emotional false memory in which (a) valence varies but arousal is held constant, (b) arousal varies but valence is held constant, or (c) arousal and valence are both varied factorially.

As yet, only one developmental study has been published with these materials, which was reported by Brainerd et al. (2010). Valence and arousal effects were separated by factorially manipulating them: Lists from each of the four Valence X Arousal cells of Table 1 were administered to samples of 7-, 11-, and 20-year-olds, followed by recognition tests for list words, critical distractors, and semantically unrelated distractors. Because response bias, as indexed by false alarms to unrelated distractors, usually declines with age (for a review, see Brainerd, Reyna, & Howe, 2009), the signal detection statistic d' was used to measure false

memory levels for each Valence X Arousal combination. The results are plotted by age in Figure 1. There are five key patterns. First and foremost, contrary to the commonsense idea that children might be more susceptible to the distortive influence of emotion than adolescents or adults, the data ran in the opposite direction. Regardless of whether arousal was higher or lower, false memory increased with age for both types of emotional valence. Second, consistent with the hypothesis that negative valence is a more salient gist than positive valence, negative valence produced higher levels of false memory than positive valence, regardless of arousal level. Third, consistent with the hypothesis that arousal interferes with verbatim retrieval's ability to suppress false memory, the higher-arousal lists produced more false memory than the lower-arousal lists, regardless of valence. Fourth, the second and third patterns waxed with development—false memory did not vary as a function of either valence or arousal in 7-year-olds. Although all of these results are of forensic interest, the fifth is perhaps of greatest significance for the reliability of witnesses' memories: The configuration of valence and arousal that produced the highest levels of false memory is the same configuration that is present in criminal cases—namely, negative valence + high arousal. Another key result that is not apparent in Figure 1 echoes some previously mentioned studies of nonemotional false memory: Emotional false memory increased so much with age that net accuracy was higher in 7-year-olds than in young adults.

Considering the legal significance of false memory for emotional experiences, it goes without saying that there is an urgent need to accelerate developmental research on this topic. The current archive of studies is so thin that research on this topic is only just beginning, and hence, it does not yet provide a firm foundation for expert scientific testimony. Also, these data only apply to circumstances in which children actually experience emotional events and therefore may be assumed to store memories of such events. If the alleged events did not happen, which is one of the two standard defense theories in criminal cases, the data are not relevant. This is not a consideration in false memory research, of course, where we have control over events, but it is a key consideration in the law, where we do not.

Testing Explanatory Principles

The studies that have been reviewed up to this point establish quite clearly that developmental reversals in false memory are robust realities. We know that they can be produced under theoretically specified conditions by tasks that range from word lists to complex thematic events to tasks that simulate important features of criminal cases, such as manipulative questioning of witnesses and eyewitness identification of suspects. On the one hand, such an existence proof of developmental reversals is significant when it comes to traditional expert testimony about the vulnerability of child witnesses' memories to falsification. In that connection, the recent tendency for some courts to rule that such testimony is not needed because such vulnerability is common knowledge evidently needs to be tempered. On the other hand, the courts require more than empirical demonstrations of age trends in false memory. Because the details of legal cases and therefore the types of events that witnesses attempt to remember vary from case to case, the most helpful scientific information in the uncertain business of judging the credibility of different sources of evidence is empirically substantiated *theoretical principles*. Such principles are especially helpful for the simple reason that they cut across different events and different contexts (Brainerd et al., 2008).

Some candidate principles, and we stress "candidate," are available in the ideas that predicted development reversals in the first place, which were rather simple, certainly simple enough for expert testimony to explain to jurors: Age increases in false memory are likely to occur in circumstances in which (a) having false memories depends on extracting

the meaning content of experience, especially connecting shared meanings across disparate events, and (b) there are obstacles to using verbatim memories of actual events to suppress false memories of related events. What sorts of findings would be needed to convert these two ideas to empirically substantiated status? Generally speaking, what is required is an archive of experimentation in which manipulations that embody the ideas are found to have two effects. First, they should drive false memory levels in theoretically predicted directions. Specifically, manipulations that make it more likely that meaning content will be processed and connected during encoding or on memory tests ought to increase false memory, while manipulations that place obstacles in the path of semantic processing or that make it more likely that verbatim traces will be used on memory tests should suppress false memories. Second, such manipulations should interact with age in predictable ways. Specifically, manipulations that enhance semantic processing should be less effective in increasing false memory in adults or adolescents than in children (because semantic processing is what children lack), and manipulations that interfere with semantic processing or shift processing should be less effective in reducing false memory in children than in adolescents or adults (because the latter age levels are possessed of superior semantic processing).

In the parlance of developmental research, findings of the first type are called *sufficiency* interactions, while findings of the second type are called *necessity* interactions. In the remainder of this section, we summarize some salient examples of manipulations that have produced each type of interaction, the ultimate consequence being to directly tie developmental reversals to the principles that originally predicted them.

Necessity Interactions

The rationale underlying necessity interactions is take-away logic; that is, the aim is to take away some process that is presumed to be crucial to age increases in false memory, and hence, the focus is on changing older subjects (who are likely to possess that process) more than younger ones. Suppose that the growth of some memory ability *M* (processing and connecting meaning in this case) is hypothesized to cause age increases in false memory. If so, manipulations that block that activity ought to reduce false memory more at ages levels where the development of *M* is reasonably far along (and false memory rates are higher) than at younger age levels, producing what are traditionally called convergent interactions in analysis of variance (see Figure 2 for illustrations). Thus, if we institute manipulations that block semantic processing in the connected-meaning paradigms (that are known to produce developmental reversals), age increases in false memory should shrink (Figure 2A), or perhaps disappear (Figure 2B), because there is more to be blocked in older than in younger subjects. The types of manipulations that have generated such interactions are ones that transparently shift subjects away from processing semantic content and forming meaning relations on connected-meaning tasks.

With the DRM illusion, categorized materials, and more complex tasks, there is an impressive collection of semantic take-away manipulations that have produced necessity interactions by either shrinking or eliminating the developmental reversal effect. The chief examples are: (a) list length (Carneiro et al., 2007; Sugrue & Hayne, 2006); (b) presenting lists pictorially (Howe, 2006; 2008); (c) presenting lists with phonological rather than semantic relations between targets (Dewhurst et al., 2011; Kahana & Cortese, 2009); (d) presenting lists with weak rather than strong semantic relations between targets (Brainerd et al., 2002); (e) measuring false memory in disabled versus nondisabled subjects (Brainerd et al., 2006; Weekes et al., 2007); (f) random versus blocked presentation of meaning-sharing events (Lampinen et al. 2006); and (g) cuing of incorrect list themes (Odegard et al., 2008). Such manipulations would be expected to partly or wholly neutralize the superior semantic processing of older subjects in one of three ways.

The first way, which is the route that has most often been taken to generate necessity interactions, is to make the semantic relations that support false memories less salient, so that older subjects are less likely to follow their normal tendency to process them. Presenting events in such a way that events share those relations do not appear together (manipulation f), presenting fewer exemplars of such relations (manipulation a), presenting events that are poor exemplars of such relations (manipulation d), presenting events in which such relations are absent (manipulation c), and cuing competing semantic relations (manipulation g) are all examples of this approach. A second way to reduce the superior semantic-processing of older subjects is to make it easier for them to rely on their superior verbatim abilities instead. Presenting events that lead to the formation of especially strong verbatim traces (manipulation b) is an example of this approach. A third way to reduce the superior semantic processing of older subjects is simply to test subjects who are known to have semantic processing limitations, so that the baseline superiority of older subjects is compromised. Manipulation e is an example of this last approach.

Some readers will be thinking that there is a more extreme class of necessity interactions—namely, Age X Manipulation cross-overs, in which connected-meaning paradigms produce developmental reversals under one level of a manipulation (e.g., the standard version of the paradigm) but produce the customary developmental decline under another level. Readers will also be thinking that the theoretical principles that predict developmental reversals anticipate that such cross-overs can occur when manipulations reduce the relevant form of semantic processing to a minimum and at the same time maximize verbatim processing. Cross-overs have indeed been detected with manipulations of that sort in experiments by Brainerd and Reyna (2007), Carneiro and Fernandez (2010), Ceci et al. (2007), and Holliday et al. (2011).

Brainerd and Reyna (2007) used a categorized list procedure that had been found to produce developmental reversals (Howe, 2006, 2007). They noted that in prior developmental work on false memory for categorized lists (Brainerd & Reyna, 2006; Brainerd, Reyna, & Kneer, 1995), false recognition of unrepresented exemplars had *decreased* with age. In that work, the study lists contained *single* exemplars of categories (e.g., false memory for *guitar* was tested following a list on which *piano* appeared), whereas in Howe's studies, the lists contained many exemplars of categories (e.g., false memory for *guitar* was tested following a list on which *piano, drums, trumpet, violin, clarinet, trombone, saxophone, and tuba* all appeared). Brainerd and Reyna (2007) proposed that this produced the contrasting age trends, in line with the hypothesis that forming meaning connections, rather than simply understanding meaning, is crucial to developmental reversals. They compared false memory for *the same* unrepresented exemplars (e.g., *guitar*) when study lists contained 8 words versus 1 words from the target category, finding as predicted that false recognition increased between age 8 and age 14 with the first procedure but decreased with the second. In the Carneiro and Fernandez (2010) study, the authors implemented a warning procedure to lower reliance on semantic processing and to increase reliance on verbatim memory in the DRM illusion. Before presenting DRM lists, children in the 4–14 range were warned that the experimenter was attempting to trick them by presenting several words that were related to each other, which would cause them to think of other related words that they would then falsely report on memory tests. The children were cautioned to be wary of this deception and to only report words that were actually presented. False recall and false recognition both increased with age in a standard (no warning) DRM condition, but decreased with age in this warning condition.

Turning to the cross-over interactions reported by Ceci et al. (2007) and Holliday et al. (2011), the methodology of the former study was discussed above—where it was noted that a developmental reversal was produced for implanted false memories using picture stories

containing objects for which (based on prior rating data) older children would be more likely than younger children to process the semantic relations that would foment false memory. However, the picture story included other objects for which (based on the same data) older children would be *less* likely to process the relevant semantic relations, eliminating their baseline superiority in semantic processing and allowing their superior verbatim memories to take over. As expected, there was an Age X Type of Picture cross-over, with these latter pictures producing developmental *decreases* in false memory. Finally, Holliday et al. generated a cross-over by comparing a standard DRM condition, in which lists are presented as whole words, to a “shallow” encoding condition in which lists were presented as word fragments with a single missing letter (e.g., *tabl*_, *si*_, *leg*_, *sea*_, and so forth for the “chair” list), and subjects were required to fill in the missing letter. This type of presentation focuses processing on words’ surface forms, rather than their meaning, and shifts performance in the direction of verbatim retrieval on memory tests (Brainerd & Reyna, 2005). Consistent with that scenario, Holliday et al. found that between the ages of 7 and 11, the false alarm rate for critical distractors increased from 31% to 43% in the standard condition, but it decreased from 26% to 14% in the shallow condition.

Sufficiency Interactions

The rationale for sufficiency interactions is giving rather than taking away; that is, the aim is to supply cognitive prosthetics for some process that is presumed to crucial to age increases in false memory, and hence, the focus is on changing younger subjects (who are unlikely to possess that process) more than younger ones. As before, suppose that the growth of some memory ability *M*, such as processing and connecting meaning, is hypothesized to cause age increases in false memory. If so, manipulations that encourage or provide support for that activity ought to enhance false memory more at younger ages levels, where the development of *M* is less far along (and false memory rates are lower), than at older age levels. This should again produce convergent interactions, but ones that are the opposite of necessity interactions (see Figure 3 for the two basic varieties). Thus, if manipulations are imposed that support semantic processing in connected-meaning paradigms that are known to produce developmental reversals, age increases in false memory should shrink (Figure 3A), or perhaps disappear (Figure 3B), because such support aids younger children more than older than older children or adults. The types of manipulations that have produced such interactions are ones that transparently encourage the processing of semantic content and the formation of meaning relations. We consider the two main examples, gist cuing and story context.

Gist cuing

An obvious example of such a manipulation, which has been found to produce both the type of sufficiency interaction in Figure 3A and the type in Figure 3B, is what Brainerd et al. (2004) called gist cuing. To date, it has been studied in experiments reported by Brainerd et al. (2004, 2006, 2008b), Holliday et al. (2008), Lampinen et al (2006), and Howe (2006). The logic of this manipulation, which derives from prior developmental studies of memory for taxonomic relations (for reviews, see Bjorklund, 1987; Bjorklund & Muir, 1988), is that younger children’s failure to form strong meaning connections among semantically related events can be ameliorated by simply telling them that they should do so—usually, by telling them to search for such relations as target materials are encoded. A general finding of those studies is that the effectiveness of gist cuing follows an inverted-U relation with age: It is most effective at intermediate age levels, such as later childhood, than during very early childhood or young adulthood. The reason, of course, as Bjorklund, Miller, Coyle, and Slawinski (1997) discussed, is that the success of gist cuing will be maximal when children have reached the age at which basic semantic processing abilities have emerged but have not

yet reached the age at which the use of those abilities is so advanced that meaning connections are formed automatically.

The initial studies of gist cuing in false memory only included two age levels and, hence, could not test for inverted-U relations. Brainerd et al. (2004) found that gist cuing elevated false memory for category exemplars more in 7-year-olds than in 14-year-olds, and Brainerd et al. (2006) found that gist cuing elevated false memory for DRM critical distractors more in 6-year-olds than in 14-year-olds. When gist cuing was compared to no-cuing conditions, both experiments produced an Age X Cuing interaction of the sort depicted in Figure 3A. In a later study, Odegard et al. 2008 studied gist cuing at 2 age levels, 11-year-olds and young adults. When a condition in which subjects were cued with the themes of DRM lists was compared to a standard presentation condition, an interaction like that in Figure 3B was obtained.

In another DRM study, Lampinen et al. (2006) were the first to test for inverted-U relations, by comparing a gist cuing condition to a standard no-cuing condition in 6-year-olds, 8-year-olds, and young adults. Their data (false alarm rates for critical distractors) are plotted in Figure 4, and comparing performance in the cuing versus no cuing conditions, two patterns emerge. First, gist cuing increased false memory more in 8-year-olds than in 6-year-olds or young adults, and, second, there was an Age X Cuing interaction of the sort depicted in Figure 2A between age 8 and young adulthood. Holliday et al. (2008) studied the effects of gist cuing on DRM false recall at five age levels: age 7, 9, 11, and 13, and 15. Like Lampinen et al., they found that cuing's tendency to increase false memory was greater at one of the intermediate age levels (13) than at earlier or later age levels.

Story context

Another, especially powerful, semantic prosthetic has been investigated in DRM studies by Dewhurst et al. (2007) and Howe and Wilkinson (2011). In this procedure, false memory in a standard presentation condition is compared to false memory in a condition in which the words on individual lists are presented in sentences that form a story that revolves around the theme of the list. To illustrate, the words for which *doctor* is the critical distractor are presented in a story about a girl named Sally who was ill, who did not take her medicine, and had to go to a clinic for treatment. The salient feature of this manipulation, of course, is that children do not have to spontaneously connect the gist among the list words because the story does it for them. When Dewhurst et al. introduced this procedure, they compared the DRM false recognition of 5-, 8-, and 11-year-olds in a story context versus a control condition. Their results are plotted in Figure 5A. In the control condition, as usual, false recognition increased with age. In the story context condition, as predicted, presenting DRM lists in stories increased false recognition far more in the youngest children than at the two older age levels. However, as can be seen in Figure 5A, there was an unanticipated pattern: Story context caused false memory to *decline* between age 5 and the two older age levels. There is no obvious explanation for this pattern, but readers may remember that the memory development literature contains occasional studies in which strategy manipulations that help younger children's memories impair older children's by interfering with the normal way that older children use well-learned strategies (see Bjorklund & Schneider, 1998; Schneider & Pressley, 1997).

However, based on a later experiment by Howe and Wilkinson (2011), the odd developmental trend in Dewhurst et al. (2007) may be a statistical aberration that does not require explanation. Howe and Wilkinson compared the DRM false recognition of 7- and 11-year-olds in a story context versus a control condition, and their results for critical distractors are plotted in Figure 5B. There, it can be seen that (a) story context greatly elevated false recognition in 7-year-olds, and (b) it also elevated false recognition in 11-

year-olds, though by only a small amount. Consequently, there was an Age X Presentation interaction of the Figure 2A variety inasmuch as developmental reversals in false memory were more marked with standard presentation than with story contexts.

Summary and Interpretive Comments

In the first section of this paper, we discussed some theoretical ideas that provide a mechanism for predicting that semantic false memory can sometimes increase between early childhood and young adulthood, especially when (a) it depends on forming meaning connections among disparate events and (b) it is difficult to use verbatim traces of those events to suppress false memories. In the second section, we exhibited many examples of studies in which this counterintuitive prediction has been confirmed and listed them in Table 1. In the present section, we took up the more fundamental question, for the law as well as for the science of memory development, of whether there is independent empirical support for the parent theoretical ideas. We have seen that there is, in the form of necessity interactions and sufficiency interactions that have been detected in several DRM studies, plus a smattering of studies with categorized materials and other paradigms.

Necessity interactions evaluate the hypothesis that if improved semantic processing, particularly improved meaning connection, is a cornerstone of age increases in semantic false memory, then it will be possible to shrink or erase those increases by placing obstacles in the path of semantic processing. We saw that there is a group of manipulations that interfere with semantic processing in various ways (e.g., by presenting fewer exemplars of a meaning or by presenting less salient meanings) that have the expected effects. Some especially powerful manipulations are able to flip-flop the relation between age and level of false memory. Sufficiency interactions evaluate the contrasting hypothesis that if improved semantic processing is a cornerstone of age increases in semantic false memory, it will also be possible to shrink or erase those increases by supplying aids to such processing. We saw that two manipulations of that sort have been investigated, gist cuing and story context, and that both have reduced age increases in false memory by increasing false memory more in younger subjects.

In short, the accumulated evidence on necessity and sufficiency interactions provides independent support for the theoretical principles that first predicted developmental reversals in false memory. It should be stressed that this does address the question of whether there are other mechanisms that also contribute to developmental reversals, and indeed, other mechanisms have been proposed (e.g., Howe, 2006; Metzger et al., 2008). Many such possibilities remain to be explored, and some will eventually be added to the list of empirically verified processes. Here, it must be remembered that what is important for the law is cumulative rather than complete knowledge. That is, what is essential is that we possess scientific understanding of some of the major factors that cause developmental reversals, not that we understand all of them. An incomplete list of validated causes can be of great assistance to triers of fact in weighing the credibility of different sources of evidence and is preferable by far to sheer ignorance.

Concluding Comments

The law has a long-standing view of the reliability of child witnesses' memories that comes down to us from the days of statutory exclusion and is grounded in much experimentation that was published in the 1980s and 1990s. According to that view, the baseline situation, which jurors must weigh in formulating verdicts, is that children's memories are less reliable than those of adolescents and adults, owing to the fact that false memory levels are higher and true memory levels are lower. That view is so ingrained that some courts have ruled that it is common knowledge on which jurors need not be educated by scientific witnesses. The

core implication for testimony is that other things being equal, when children's evidence conflicts with that of older witnesses', as it very often does, children's recountings are more likely to be the erroneous ones. It is on this specific point that the expanding literature on developmental reversals in false memory makes a major contribution by giving new clarity to "other things being equal."

Theoretical considerations predict and data confirm that there are classes of events—specifically, ones that are connected by salient semantic relations—for which other things are not equal because older subjects are more likely to process and store the information that generates false memories than children are. Research with various paradigms in which false memories are rooted in meaning connections has revealed the predicted age increases in errors, and in paradigms that produce especially large increases, net accuracy has sometimes declined with age. With those same tasks, the memory processes that were used to predict these patterns in the first place have been found to be both necessary and sufficient to produce them. A feature of these results that is of overriding legal significance is that, contrary to the traditional view of how witness reliability varies with age, they provide grounds for supposing that age increases in false memory for real-world events may be quite wide spread. That is because connected-meaning paradigms tap one of the most pervasive features of real-world experience—namely, that distinct events share meaning. Our everyday lives consist in large measure of meaningful clusters of events that occur in script-like situations, so that events that share the salient meanings of the index situations: arising and getting ready for the day, preparing and eating breakfast, driving to work or school, attending class, eating lunch with friends, shopping at a mall or grocery store, attending a movie or concert, and so on and on. Thus, daily experience is dominated by events that exhibit the very properties that theory says (and data confirm) foment age increases in false memory and produce net reductions in accuracy.

A related point of legal significance is that developmental increases and decreases in false memory *are both lawful patterns*. Some might be tempted to think that as opposite age trends have been identified, developmental findings on false memory are just a welter of inconsistency; that nothing can be concluded that is substantial enough to meet a Frye or Daubert test in the courtroom. Such thinking would be a capital mistake because we saw at the outset that (a) opposite developmental trends were *predicted* on theoretical grounds before they were detected, and (b) the same theoretical ideas that predicted opposite trends specified the exact conditions under which each age trend will be observed. Thus, the actual conclusion for the law is that developmental findings on false memory provide firm empirical support for theoretical principles that form the basis of expert testimony.

We stress that although accumulated research on developmental reversals has fundamental implications for the law's view of the reliability of child witnesses, it should be seen as an edifice whose construction is just beginning. Far more remains to be done than what has already been accomplished. So far, research has put a pair of forensic cornerstones in place: The notion that children's evidence is necessarily more infected with false memories than adolescents' or adults' is mistaken, and there are domains of everyday experience in which age increases in false memory are likely to occur and may even be the rule. However, a good deal of theoretical analysis and focused experimentation are needed to deepen the forensic connections of such research; to identify forensic situations and forensically relevant variables that are reliable sources of developmental reversals. We close with a prominent example of one such situation, memory suggestion, and one such variable, emotional memory, and also with a call for additional findings about the adolescent age range.

Memory Suggestion

From the first, developmental research on the reliability of child witnesses' evidence has been dominated by studies of suggestibility, using the Loftus misinformation paradigm (Bruck & Ceci, 1999; Ceci & Bruck, 1993). That is because a prominent way in which children witnesses' memories are contaminated is via suggestive questioning by investigators, parents, teachers, therapists, friends, and so on (Ceci & Bruck, 1995). The literature now contains dozens of studies in which false memory for suggested events was found to decline between early childhood and young adulthood, so that expert witnesses have naturally testified that childhood is an especially suggestion-prone time of life (Bruck & Ceci, 1999). Against that backdrop, there are newer studies in which the misinformation paradigm has been modified to implement theoretical principles that predict developmental reversals, and false memories have increased rather than decreased with age. However, that group consists of only five recent studies (Ceci et al., 2007; Connolly & Price, 2006; Fazio & Marsh, 2008; Principe et al., 2008; Ross et al., 2006), plus two earlier ones by Pezdek and Rowe (1995, 1997). It is true that the non-accidental nature of the recent studies—that they produced age increases under theoretically specified conditions—counts heavily in their favor. Nevertheless, seven studies that have produced age increases seems a paltry sum in comparison to the dozens that have produced age declines (in many countries and multiple languages).

Consequently, it is natural for the law to take a wait-and-see approach—adopting the working hypothesis that until the data become overwhelming, the studies that have produced developmental reversals could be statistical aberrations. That, in turn, means continuing to rely on the age decrease pattern when assessing the reliability of children's evidence. In fact, there are signs of such attitudes in recent discussions of child witnesses in the psycho-legal literature (e.g., Bull, 2010; Neoh, & Mellor, 2009), discussions that reaffirm the age decrease pattern and make no mention of studies that generated the opposite pattern. The message could not be clearer, we think: To ensure that developmental reversals in suggestibility penetrate the forensic literature and expert testimony, more experimentation that demonstrates the reality of those reversals and establishes the conditions that produce them is an urgent priority.

Emotional False Memory

A central feature of crimes is that the events are emotional—so that the scientific study of testimony must confront the question of whether emotional content is distortive or protective of the accuracy of witnesses' recountings and, with child witnesses, whether such effects vary with age. We saw that methodological tools are available to evaluate such predictions, in the form of emotion-induction tasks (e.g., Storbeck & Clore, 2005) and emotional connected-meaning tasks (e.g., Budson et al., 2006). Considering the simplicity of these methods and the fact that they do not raise ethical concerns, it is remarkable that so few developmental studies implementing them have been published. Although those studies are few in number, they have already generated data of potential significance for the law.

Thus, a data archive on the development of emotional false memory is beginning to accumulate, but the operative word is "beginning." As is often true in the early phases of research, the extant literature is replete with empirical inconsistencies that will need to be sorted out and with empirical gaps that need to be filled in. With respect to inconsistencies, perhaps the most obvious one revolves around the two core methods of testing episodic memory, recognition and recall. The question of whether recall or recognition is more apt to stimulate false memories during criminal investigation has long figured prominently in research, especially in research on best-practice protocols for investigative interviewing of

children (Ceci & Bruck, 1995; Poole & Lamb, 1998). It is particularly important to answer this question for emotional false memory, but so far, every possible configuration of results has been reported. Howe and associates (2007; Howe et al., 2010) initially found that levels of false memory were higher for emotional than neutral materials with recognition, with the reverse being true for recall. Goodman et al. (2011) then found that emotional false memory was higher than neutral false memory with recall, but the reverse was true for recognition. Very recently, Otgaar et al. (2012) found that recall tests produced equivalent levels of emotional and neutral false memory. To say the least, inconsistent findings as to how emotional false memory is influenced by recognition versus recall do not provide a firm basis for best-practice recommendations about forensic interviewing or for expert testimony.

With respect to gaps in research, there is a yawning chasm when it comes to methodological breadth. All the studies that we reviewed relied on procedures in which emotional content was embedded in the to-be-remembered events, and all but one (Fernandez-Dols et al., 2008) relied on some version of the emotional DRM task. One of the two core approaches to studying emotional false memory, mood induction, does not seem to have been used in any published developmental research. This circumstance, too, is not a firm basis for forensic interviewing or expert testimony because there are no guarantees that the findings that we reviewed for the first approach to emotional false memory will hold for the second. On the contrary, the second approach has been implemented in recent studies of emotional false memory *in adults* to test predictions that are the *opposite* of some of those findings. For instance, as mentioned earlier, Storbeck and Clore (2005) proposed that if negative moods are induced in subjects before they encode to-be-remembered information, their processing of that information will be shifted toward surface details, relative to neutral or positive moods, simultaneously strengthening the verbatim traces that suppress false memories and weakening the gist traces that foment false memories. In their research, Storbeck and Clore found, as predicted, that false memory was lowest in subjects who had been placed in negative moods. Later, Corson and Verrier (2007) replicated this pattern and showed that it was chiefly due to the arousal rather than the valence component of negative moods. If being placed in negative moods lowers false memory in adults, age increases in false memory ought to be *smaller* for subjects who were in negative moods than for subjects who were in neutral or positive ones.

In short, developmental research on emotional false memory has produced some tantalizing results. However, research on this topic is far away from achieving the levels of empirical agreement and methodological breadth that are commensurate with the importance of emotional memory in legal cases.

Adolescence

Another important lacuna in the data base is the general absence of findings on adolescence, particularly middle-to-late adolescence. In Table 1, it can be seen that in 23 articles, the age spread of the subject samples ranges from childhood to adulthood. However, subjects from the adolescent years were included in very few of those studies. In the typical study (e.g., Howe, 2005; Lampinen et al., 2006), the comparison has simply been between children and young adults, with adolescents not being tested. In the few studies in which adolescents were included (e.g., Brainerd et al., 20002), only younger adolescents were tested. Thus, findings on variables that contribute to age increases in false memory throughout the course of adolescence years are extremely scarce.

Epilogue

Developmental research on false memory has provided what, when the history of this era is written, it will surely be seen as a singularly productive collaboration between the law and

psychological science inasmuch as such research has truly changed how cases are investigated and tried. It has also altered the content of expert scientific testimony and what the courts view as acceptable scientific testimony about memory. More changes, it would seem, are in the offing that reflect a new generation of findings about developmental reversals in false memory, findings that challenge accepted wisdom about the reliability of children's memories.

Acknowledgments

Portions of the research were supported by National Institutes of Health Grant 1RC1AG036915-01.

References

- Ackerman BP. The sources of children's source errors in judging causal inferences. *Journal of Experimental Child Psychology*. 1992; 54:90–119.
- Ackerman BP. Children's source errors in referential communication. *Journal of Experimental Child Psychology*. 1994; 58:432–464.
- Ackil JK, Zaragoza MS. Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology*. 1995; 60:57–83. [PubMed: 7666038]
- Ackil JK, Zaragoza MS. Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology*. 1998; 34:1358–1372. [PubMed: 9823517]
- Anastasi JS, Lewis S, Quinlan F. Examining differences in the levels of false memories in children and adults using child-normed lists. *Developmental Psychology*. 2008; 44:889–894. [PubMed: 18473653]
- Arndt J, Reder LM. The effect of distinctive visual information on false recognition. *Journal of Memory & Language*. 2003; 48:1–15. [PubMed: 19079562]
- Banaji MR, Crowder RG. The bankruptcy of everyday memory. *American Psychologist*. 1989; 44:1185–1193.
- Binet, A. *La suggestibilité*. Paris: Schleicher Feres; 1900.
- Bjorklund DF. How changes in knowledge base contribute to the development of children's memory. *Developmental Review*. 1987; 7:93–130.
- Bjorklund, DF. *Children's thinking: Cognitive development and individual differences*. Belmont, CA: Wadsworth; 2004.
- Bjorklund DF, Bjorklund BR, Brown RD. Children's susceptibility to repeated questions: How misinformation changes children's answers and their minds. *Applied Developmental Science*. 1998; 2:99–111.
- Bjorklund DF, Cassel WS, Bjorklund BR, Brown RD, Park CL, Ernst K, Owen FA. 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*. 2000; 14:421–433.
- Bjorklund DF, Hock HH. Age differences in the temporal locus of memory organization in children's recall. *Journal of Experimental Child Psychology*. 1982; 33:347–362.
- Bjorklund DF, Jacobs JW. Associative and categorical processes in children's memory: The role of automaticity in the development of free recall. *Journal of Experimental Child Psychology*. 1985; 39:599–617.
- Bjorklund DF, Miller PH, Coyle TR, Slawinski JL. Instructing children to use memory strategies: Evidence of utilization deficiencies in memory training studies. *Developmental Review*. 1997; 17:411–441.
- Bjorklund DF, Muir JE. Children's development of free recall memory: Remembering on their own. *Annals of Child Development*. 1988; 5:79–123.
- Bouwmeester S, Verkoeijen PPJL. Latent variable modeling of cognitive processes in true and false. *Journal of Experimental Psychology: General*. 2010; 139:365–381. [PubMed: 20438256]

- Bouwmeester S, Vermunt JK, Sijtsma K. Development and individual differences in transitive reasoning: A fuzzy trace theory approach. *Developmental Review*. 2007; 27:41–74.
- Bradley, MM.; Lang, PJ. Technical Report C-1, Center for Research in Psychophysiology. University of Florida; 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings.
- Brainerd CJ, Forrest TJ, Karibian D, Reyna VF. Development of the false-memory illusion. *Developmental Psychology*. 2006; 42:662–679.
- Brainerd CJ, Gordon LL. Development of verbatim and gist memory for numbers. *Developmental Psychology*. 1994; 30:163–177.
- Brainerd CJ, Holliday RE, Reyna VF. Behavioral measurement of remembering phenomenologies: So simple a child can do it. *Child Development*. 2004; 75:505–522. [PubMed: 15056202]
- Brainerd CJ, Holliday RE, Reyna VF, Yang Y, Toglia MP. Developmental reversals in false memory: effects of emotional valence and arousal. *Journal of Experimental Child Psychology*. 2010; 107:137–154. [PubMed: 20547393]
- Brainerd CJ, Mojardin AH. Children's spontaneous false memories for narrative statements: Long-term persistence and mere testing effects. *Child Development*. 1998; 69:1361–1377. [PubMed: 9839421]
- Brainerd CJ, Reyna VF. Explaining memory free reasoning. *Psychological Science*. 1992; 3:332–339.
- Brainerd CJ, Reyna VF. Autosuggestibility in memory development. *Cognitive Psychology*. 1995; 28:65–101. [PubMed: 7895469]
- Brainerd CJ, Reyna VF. Mere memory testing creates false memories in children. *Developmental Psychology*. 1996; 32:467–476.
- Brainerd CJ, Reyna VF. Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology*. 1998; 71:81–129. [PubMed: 9843617]
- Brainerd CJ, Reyna VF. Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*. 2002; 11:164–169.
- Brainerd CJ, Reyna VF. Fuzzy-trace theory and memory development. *Developmental Review*. 2004; 24:396–439.
- Brainerd, CJ.; Reyna, VF. The science of false memory. New York: Oxford University Press; 2005.
- Brainerd CJ, Reyna VF. Explaining developmental reversals in false memory. *Psychological Science*. 2007; 18:442–448. [PubMed: 17576285]
- Brainerd CJ, Reyna VF, Brandse E. Are children's false memories more persistent than their true memories? *Psychological Science*. 1995; 6:359–364.
- Brainerd CJ, Reyna VF, Ceci SJ. Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin*. 2008a; 134:343–382. [PubMed: 18444700]
- Brainerd CJ, Reyna VF, Ceci SJ, Holliday RE. Understanding developmental reversals in false memory. *Psychological Bulletin*. 2008b; 134:773–777.
- Brainerd CJ, Reyna VF, Estrada S. Recollection rejection of false narrative statements. *Memory*. 2006; 14:672–691. [PubMed: 16829486]
- Brainerd CJ, Reyna VF, Forrest TJ. Are young children susceptible to the false-memory illusion? *Child Development*. 2002; 73:1363–1377. [PubMed: 12361306]
- Brainerd CJ, Reyna VF, Howe ML. Trichotomous processes in early memory development, aging, and neurocognitive impairment: A unified theory. *Psychological Review*. 2009; 116:732–783.
- Brainerd CJ, Reyna VF, Kneer R. False-recognition reversal: When similarity is distinctive. *Journal of Memory and Language*. 1995; 34:157–185.
- Brainerd CJ, Reyna VF, Wright R, Mojardin AH. Recollection rejection: False-memory editing in children and adults. *Psychological Review*. 2003; 110:762–784. [PubMed: 14599242]
- Brainerd CJ, Stein LM, Silveira RA, Rohenkohl G, Reyna VF. How does negative emotion cause false memories? *Psychological Science*. 2008c; 19:919–925. [PubMed: 18947358]
- Brainerd CJ, Wright R, Reyna VF, Payne DG. Dual-retrieval processes in recall. *Journal of Memory and Language*. 2002; 46:120–152.
- Brainerd CJ, Yang Y, Howe ML, Reyna VF, Mills BA. Semantic processing in “associative” false memory. *Psychonomic Bulletin & Review*. 2008b; 15:1035–1053. [PubMed: 19001566]

- Brown SC, Flores LM, Goodman KM, Conover JN. Clustering and recall: Do high clusterers recall more than low clusterers because of clustering? *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1991; 17:710–721.
- Bruck M, Ceci SJ. The suggestibility of children's memory. *Annual Review of Psychology*. 1999; 50:419–439.
- Bruck M, Ceci SJ, Francoeur E, Barr R. I hardly cried when I got my shot: Influencing children's reports about a visit to their pediatrician. *Child Development*. 1995; 66:193–208. [PubMed: 7497825]
- Budson AE, Todman RW, Chong H, Adams EH, Kensinger EA, Krangel TS. False recognition of emotional word lists in aging and Alzheimer's disease. *Cognitive and Behavioral Neurology*. 2006; 19:71–78. [PubMed: 16783129]
- Bull R. The investigative interviewing of children and other vulnerable witnesses: Psychological research and working/professional practice. *Legal and Criminological Psychology*. 2010; 15:5–23.
- Carneiro P, Albuquerque P, Fernandez A. Opposite developmental trends for false recognition of basic and superordinate names. *Memory*. 2009a; 17:411–427. [PubMed: 19241218]
- Carneiro P, Albuquerque P, Fernandez A, Esteves F. Analyzing false memories in children with associative lists specific for their age. *Child Development*. 2007; 78:1171–1185. [PubMed: 17650132]
- Carneiro P, Fernandez A. Age differences in the rejection of false memories: The effects of giving warning instructions and slowing the presentation rate. *Journal of Experimental Child Psychology*. 2010; 105:81–97. [PubMed: 19875129]
- Carneiro P, Fernandez A, Dias AR. The influence of theme identifiability on false memories: Evidence for age-dependent opposite effects. *Memory & Cognition*. 2009b; 37:115–129.
- Carstensen LL, Mikels JA. At the intersection of emotion and cognition: Aging and the positivity effect. *Current Directions in Psychological Science*. 2005; 14:117–121.
- Ceci SJ, Bruck M. The suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*. 1993; 113:403–439. [PubMed: 8316609]
- Ceci, SJ.; Bruck, M. *Jeopardy in the courtroom*. Washington, DC: American Psychological Association; 1995.
- Ceci SJ, Bruck M. The ontogeny and durability of true and false memories: A dual trace account. *Journal of Experimental Child Psychology*. 1998; 71:165–169. [PubMed: 9843621]
- Ceci SJ, Friedman RD. The suggestibility of children: Scientific research and legal implications. *Cornell Law Review*. 2000; 86:34–108.
- Ceci SJ, Papierno PD, Kulkofsky S. Representational constraints on children's suggestibility. *Psychological Science*. 2007; 18:503–509. [PubMed: 17576263]
- Ceci SJ, Ross DF, Toglia MP. Suggestibility in children's memory: Psycholegal implications. *Journal of Experimental Psychology: General*. 1987; 116:38–49.
- Chi, MTH.; Ceci, SJ. Content knowledge: Its representation and restructuring in memory development. In: Reese, HW.; Lipsett, L., editors. *Advances in Child Development and Behavior*. Vol. 20. 1987. p. 91-146.
- The Committee on the Judiciary of the House of Representatives. *Federal rules of evidence*. Washington, DC: United States Government Printing Office; 2006.
- Connolly DA, Price HL. 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*. 2006; 93:207–223. [PubMed: 16111696]
- Corson Y, Verrier N. Emotions and false memories: Valence or arousal? *Psychological Science*. 2007; 18:208–211. [PubMed: 17444912]
- Deese J. On the prediction of occurrence of certain verbal intrusions in free recall. *Journal of Experimental Psychology*. 1959; 58:17–22. [PubMed: 13664879]
- Dewhurst SA, Howe ML, Berry DM, Knott LM. Test-induced priming increases false recognition in older but not younger children. *Journal of Experimental Child Psychology*. 2011; 111:101–107. [PubMed: 21908005]

- Dewhurst SA, Pursglove RC, Lewis C. Story contexts increase susceptibility to the DRM illusion in 5-year-olds. *Developmental Science*. 2007; 10:374–378. [PubMed: 17444977]
- Dewhurst SA, Robinson CA. False memories in children: Evidence for a shift from phonological to semantic associations. *Psychological Science*. 2004; 15:782–786. [PubMed: 15482451]
- Fazio LK, Marsh EJ. Older, not younger, children learn more false facts from stories. *Cognition*. 2008; 106:1081–1089. [PubMed: 17540354]
- Fernandez-Dols JM, Carrera P, Barchard KA, Gacitua M. False recognition of facial expressions of emotion: Causes and implications. *Emotion*. 2008; 8:530–539. [PubMed: 18729583]
- Fisher AV, Sloutsky VM. When induction meets memory: Evidence for a gradual transition from similarity-based to category-based induction. *Child Development*. 2005; 76:583–597. [PubMed: 15892780]
- Gallo DA. Associative illusions of memory: False memory research in DRM and related tasks. New York: Psychology Press; 2006.
- Gallo D. False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*. 2010; 7:833–848.
- Garcia-Marques L, Ferreira M, Nunes L, Garrido MV, Garcia-Marques T. False Memories and Impressions of Personality. *Social Cognition*. 2010; 28:556–568.
- Gernsbacher MA. Surface information loss in comprehension. *Cognitive Psychology*. 1985; 17:324–363.
- Ghetti S, Qin J, Goodman GS. False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology*. 2002; 38:705–718. [PubMed: 12220049]
- Goodman GS. Children's eyewitness memory: A modern history and contemporary commentary. *Journal of Social Issues*. 2006; 62(4):811–832.
- Goodman GS, Ogle CM, Block SD, Harris LS, Larson RP, Augusti EM, Urquiza A. False memory for trauma-related Deese-Roediger-McDermott lists in adolescents and adults with histories of child sexual abuse. *Development and Psychopathology*. 2011; 23:71–86.
- Goodman GS, Quas JA, Batterman-Faunce JM, Riddlesberger MM, Kuhn J. Predictors of accurate and inaccurate memories of traumatic events experienced in childhood. *Consciousness and Cognition*. 1994; 3:269–294.
- Goodman GS, Schaaf JM. Over a decade of research on children's eyewitness testimony: What have we learned? Where do we go from here? *Applied Cognitive Psychology*. 1997; 11:S5–S20.
- Holliday RE, Brainerd CJ, Reyna VF. Recall of details never experienced: Effects of age, repetition, and semantic cues. *Cognitive Development*. 2008; 23:67–78.
- Holliday RE, Brainerd CJ, Reyna VF. Developmental reversals in false memory: Now you see them, now you don't. *Developmental Psychology*. 2011; 47:442–449. [PubMed: 21142371]
- Holliday RE, Reyna VF, Hayes BK. Memory processes underlying misinformation effects in child witnesses. *Developmental Review*. 2002; 22:37–77.
- Holliday RE, Weekes BS. Dissociated developmental trajectories for semantic and phonological false memories. *Memory*. 2006; 14:264–636.
- Howe ML. Misleading children's story recall: Forgetting and reminiscence of the facts. *Developmental Psychology*. 1991; 27:746–762.
- Howe ML. Children (but not adults) can inhibit false memories. *Psychological Science*. 2005; 16:927–931. [PubMed: 16313654]
- Howe ML. Developmentally invariant dissociations in children's true and false memories: Not all relatedness is created equal. *Child Development*. 2006; 77:1112–1123. [PubMed: 16942509]
- Howe ML. Children's emotional false memories. *Psychological Science*. 2007; 18:856–860. [PubMed: 17894601]
- Howe ML. Visual distinctiveness and the development of children's false memories. *Child Development*. 2008; 79:65–79. [PubMed: 18269509]
- Howe ML, Candel I, Otgaar H, Malone C, Wimmer MC. Valence and the development of immediate and long-term false memory illusions. *Memory*. 2010; 18:58–75. [PubMed: 20391177]
- Howe ML, Cicchetti D, Toth SL, Cerrito BM. True and false memories in maltreated children. *Child Development*. 2004; 75:1402–1417. [PubMed: 15369522]

- Howe ML, Gagnon N, Thouas L. Development of false memories in bilingual children and adults. *Journal of Memory and Language*. 2007; 58:669–681.
- Howe ML, Wimmer MC, Gagnon N, Plumpton S. An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language*. 2009; 60:229–251.
- Howe ML, Wilkinson S. Using story contexts to bias children's true and false memories. *Journal of Experimental Child Psychology*. 2011; 108:75–95.
- Isen, A. Positive affect, cognitive processes, and social behavior. In: Berkowitz, L., editor. *Advances in experimental social psychology*. Vol. 20. San Diego: Academic Press; 1987. p. 203–253.
- Israel L, Schacter DL. Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review*. 1997; 4:577–581.
- Kassin SM, Kiechel KL. The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science*. 1996; 7:125–128.
- Kensinger EA. Remembering and emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*. 2004; 15:241–251. [PubMed: 15526549]
- Khanna MM, Cortese MJ. Children and adults are differentially affected by presentation modality in the DRM paradigm. *Applied Cognitive Psychology*. 2009; 23:859–877.
- Kintsch W, Welsch D, Schmalhofer F, Zimny S. Sentence memory: A theoretical analysis. *Journal of Memory & Language*. 1990; 29:133–159.
- Knott LM, Howe ML, Wimmer MC, Dewhurst SA. The development of automatic and controlled inhibitory retrieval processes in true and false recall. *Journal of Experimental Child Psychology*. 2011; 109:91–108. [PubMed: 21320706]
- Koriat A, Goldsmith M. Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*. 1994; 123:297–315. [PubMed: 7931094]
- Koriat A, Goldsmith M. Monitoring and control processes in the strategic regulation of memory. *Psychological Review*. 1996; 103:490–417. [PubMed: 8759045]
- Lamb ME, Orbach Y, Hershkowitz I, Esplin PW, Horowitz D. A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD investigating interview protocol. *Child Abuse and Neglect*. 2007; 31:1201–1231. [PubMed: 18023872]
- Lampinen JM, Copeland SM, Neuschatz JS. Recollections of things schematic: Room schemas revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2001; 27:1211–1222.
- Lampinen JM, Leding JK, Reed KB, Odegard TN. Global gist extraction in children and adults. *Memory*. 2006; 14:952–964. [PubMed: 17077030]
- Lang, PJ.; Bradley, MM.; Cuthbert, BN. Technical Report A-6. Gainesville, FL: University of Florida; 2005. International affective picture system (IAPS): Affective ratings of pictures and instruction manual.
- Loftus EF. Leading questions and eyewitness report. *Cognitive Psychology*. 1975; 7:560–572.
- Loftus, EF.; Ketcham, K. *The myth of repressed memory*. New York: St. Martin's Press; 1994.
- Loftus EF, Miller DG, Burns HJ. Semantic integration of verbal information into visual memory. *Journal of Experimental Psychology: Human Learning and Memory*. 1978; 4:19–31. [PubMed: 621467]
- Lyons KE, Gheiti S, Cornoldi C. Age differences in the contribution of recollection and familiarity to false-memory formation: a new paradigm to examine developmental reversals. *Developmental Science*. 2010; 13:355–362. [PubMed: 20136932]
- Marin BV, Holmes DL, Guth M, Kovac P. The potential of children as eyewitnesses. *Law and Human Behavior*. 1979; 3:295–305.
- McAuliff BD, Kovera MB. Estimating the effects of misleading information on witness accuracy: Can experts tell jurors something they don't already know? *Applied Cognitive Psychology*. 2007; 21:849–870.

- McAuliff, BD.; Nicholson, E.; Ravanshenas, D. 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. 2007 Mar.
- McGough, LS. Child witnesses: Fragile voices in the American legal system. New Haven, CT: Yale University Press; 1993.
- Metzger RL, Warren AR, Price JD, Reed AW, Shelton J, Williams D. Do children "DRM" like adults? False memory production in children. *Developmental Psychology*. 2008; 44:169–181. [PubMed: 18194015]
- Nelson, DL.; McEvoy, CL.; Schreiber, TA. The University of South Florida word association, rhyme, and word fragment norms. Tampa: University of South Florida; 1999. Unpublished manuscript
- Neoh J, Mellor D. Professional issues related to allegations and assessment of child sexual abuse in the context of family court litigation. *Psychiatry, Psychology, and Law*. 2009; 16:303–321.
- Odegard TN, Cooper CM, Lampinen JM, Reyna VF, Brainerd CJ. Children's eyewitness memory for multiple real-life events. *Child Development*. 2009; 80:1877–1890. [PubMed: 19930357]
- Odegard TN, Holliday RE, Brainerd CJ, Reyna VF. Attention to global-gist processing eliminates age effects in false memories. *Journal of Experimental Child Psychology*. 2008; 99:96–113. [PubMed: 17950306]
- Otgaar H, Peters M, Howe ML. Dividing attention lowers children's but increases adults' false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2012; 38:204–210.
- Otgaar H, Smeets T. Adaptive memory: Survival processing increases both true and false memory in adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2010; 36:1010–1016.
- Otgaar H, Verschuere B, Meijer EH, van Oorsouw K. The origin of children's implanted false memories: Memory traces or compliance. *Acta Psychologica*. (in press).
- Payne DG, Elie CJ, Blackwell JM, Neuschatz JS. Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory & Language*. 1996; 35:261–285.
- Paz-Alonso PM, Gheiti S, Donohue SE, Goodman GS, Bunge SA. Neurodevelopmental correlates of true and false recognition. *Cerebral Cortex*. 2008; 18:2208–2216. [PubMed: 18203693]
- Petersen RC. Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*. 2004; 256:183–194. [PubMed: 15324362]
- Pezdek K, Roe C. The effect of memory trace strength on suggestibility. *Journal of Experimental Child Psychology*. 1995; 60:116–128.
- Pezdek K, Roe C. The suggestibility of children's memory for being touched: Planting, erasing, and changing memories. *Law and Human Behavior*. 1997; 21:95–106. [PubMed: 9058575]
- Pipe ME, Gee S, Wilson JC, Egerton JM. Children's recall 1 or 2 years after an event. *Developmental Psychology*. 1999; 35:781–789. [PubMed: 10380868]
- Poole, DA.; Lamb, ME. Investigative interviews of children. Washington, DC: American Psychological Association; 1998.
- Poole DA, Lindsay DS. Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology*. 1995; 60:129–154.
- Poole DA, Lindsay DS. Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied*. 2001; 7:27–50. [PubMed: 11577617]
- Poole DA, Lindsay DS. Reducing child witnesses' false reports of misinformation from parents. *Journal of Experimental Child Psychology*. 2002; 81:117–140. [PubMed: 11786006]
- Poole DA, White LT. Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*. 1991; 27:975–986.
- Principe GF, Guiliano S, Root C. Rumor mongering and remembering: How rumors originating in children's inferences can affect memory. *Journal of Experimental Child Psychology*. 2008; 99:135–155. [PubMed: 18155719]
- Quas JA, Qin JJ, Schaaf JM, Goodman GS. Individual differences in children's and adults' suggestibility and false event memory. *Learning and Individual Differences*. 1997; 9:359–390.

- Reyna VF. Conceptions of memory development, with implications for reasoning and decision making. *Annals of Child Development*. 1996; 12:87–118.
- Reyna VF, Brainerd CJ. Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*. 1995; 7:1–75.
- Reyna VF, Ellis SC. Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological Science*. 1994; 5:275–279.
- Reyna VF, Holliday R, Marche T. Explaining the development of false memories. *Developmental Review*. 2002; 22:436–489.
- Reyna VF, Kiernan B. Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology*. 1994; 30:178–191.
- Reyna VF, Kiernan B. Children's memory and interpretation of psychological metaphors. *Metaphor and Symbolic Activity*. 1995; 10:309–331.
- Reyna VF, Lloyd F. Theories of false memory in children and adults. *Learning and Individual Differences*. 1997; 9:95–123.
- Reyna, VF.; Mills, B.; Estrada, S.; Brainerd, CJ. False memory in children: Data, theory, and legal implications. In: Toglia, MP.; Read, JD.; Ross, DF.; Lindsay, RCL., editors. *Handbook of eyewitness psychology*. Mahwah, NJ: Erlbaum; 2007. p. 479-507.
- Rivers SE, Reyna VF, Mills B. Risk taking under the influence: A fuzzy-trace theory of emotion in adolescence. *Developmental Review*. 2008; 28:107–144. [PubMed: 19255597]
- Roediger HL III, McDermott KB. Creating false memories: Remembering words not presented on lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1995; 21:803–814.
- Roediger HL III, Watson JM, McDermott KB, Gallo DA. Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*. 2001; 8:385–405. [PubMed: 11700893]
- Rohrer D, Pashler H. Recent research on human learning challenges conventional instructional strategies. *Educational Researcher*. 2010; 39:406–412.
- Ross DF, Marsil DF, Benton TR, Hoffman R, Warren AR, Lindsay RCL, Metzger R. Children's susceptibility to misidentifying a familiar bystander from a lineup: When younger is better. *Law and Human Behavior*. 2006; 30:249–257. [PubMed: 16729204]
- Schneider, W.; Bjorklund, DF. Memory. In: Damon, W.; Kuhn, D.; Siegler, RS., editors. *Handbook of child psychology: Vol. 2. Cognition, perception, and language*. 5th ed.. New York: Wiley; 1998. p. 467-522.
- Schneider, W.; Pressley, M. *Memory development between two and twenty*. Englewood Cliffs, NJ: Erlbaum; 1997.
- Seamon JG, Luo CR, Kopecky JJ, Price CA, Rothschild L, Fung NS, Schwartz MA. Are false memories more difficult to forget than accurate memories?: The effect of retention interval on recall and recognition. *Memory & Cognition*. 2002; 30:1054–1064.
- Sloutsky VM, Fisher AV. When development and learning decrease memory. *Psychological Science*. 2004; 15:553–558. [PubMed: 15271001]
- Small WS. Suggestibility of children, *Pedagogical Seminary*. 1896; 13:176.
- Sommers MS, Lewis BP. Who really lives next door: creating false memories with phonological neighbors. *Journal of Memory and Language*. 1999; 40:83–108.
- Stern W. Abstracts of lectures on the psychology of testimony and the study of individuality. *American Journal of Psychology*. 1910; 21:270–282.
- Storbeck J, Clore GL. With sadness comes accuracy; with happiness, false memory: Mood and the false memory effect. *Psychological Science*. 2005; 16:785–791. [PubMed: 16181441]
- Sugrue K, Hayne H. False memories produced by children and adults in the DRM paradigm. *Applied Cognitive Psychology*. 2006; 20:625–631.
- Sugrue K, Strange D, Hayne H. False memories in the DRM paradigm age-related differences in lure activation and source monitoring. *Experimental Psychology*. 2009; 56:354–360. [PubMed: 19447751]
- Toglia, MP.; Battig, WF. *Handbook of semantic word norms*. Hillsdale, NJ: Erlbaum; 1978.

- Tulving E, Thomson DM. Retrieval processes in recognition memory: Effects of associative context. *Journal of Experimental Psychology*. 1971; 87:116–124.
- Verkoeijen PPJL, Bouwmeester S. Gist processing in free recall and recognition: Latent variable modeling of children's and adults' true and false memories. *Journal of Cognitive Psychology*. (in press).
- Weekes BS, Hamilton S, Oakhill J, Holliday RE. False recollection in children with reading comprehension difficulties. *Cognition*. 2007; 106:222–233. [PubMed: 17349990]
- Wells GL, Small M, Penrod S, Malpass RS, Fulero SM, Brimacombe CAE. Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*. 1998; 23:603–648.
- Whipple GM. The observer as reporter: A survey of the psychology of testimony. *Psychological Bulletin*. 1909; 6:153–170.
- Wilburn C, Feeney A. Do development and learning really decrease memory? On similarity and category-based induction in adults and children. *Cognition*. 2008; 106:1451–1464. [PubMed: 17604015]
- Wimmer MC, Howe ML. Are children's memory illusions created differently from those of adults? Evidence from levels-of-processing and divided attention paradigms. *Journal of Experimental Child Psychology*. 2011; 107:31–49. [PubMed: 20417937]

Highlights

- We review developmental studies of children's false memory.
- The studies show that false memories can increase as well as decrease with age.
- The conditions under which each trend occurs are theoretically well specified.

Figure 1. Developmental increases in negatively-valenced false memories (high and low arousal) and positively-valenced false memories (high and low arousal) from an experiment reported by Brainerd, Holliday, Reyna, Yang, and Toglia (2010).

Figure 2.

Basic types of necessity interactions in semantic take-away studies of developmental reversals. Panel A = manipulations that shrink developmental reversals, Panel B = manipulations that eliminate developmental reversals, and Panel C = manipulations that produce cross overs.

Figure 3.

Basic types of necessity interactions in semantic prosthetic studies of developmental reversals. Panel A = manipulations that shrink developmental reversals, and Panel B = manipulations that eliminate developmental reversals.

Figure 4. Effects of gist cuing versus standard presentation on DRM false memory at three age levels. The plotted data are from Lampinen, Leding, Reed, and Odegard (2006).

Figure 5.

Effects of story context presentation versus standard presentation on DRM false memory at different age levels. The data in Panel A are from Dewhurst, Pursglove, and Lewis (2007).. The data in Panel B are from Howe and Wilkinson (2011).

Table 1**Recent Studies of Developmental Reversals in False Memory**

Articles	Age span	# Exps. reported in article	Memory tests	Key Results
Anastasi et al. (2008)	5-adult	1	DRM lists	Age increases in false recognition and false recall for both child- and adult-normed DRM lists.
Bouwmeester & Verkoeijen (2010)	7–12	1	Dutch DRM lists	Age increases in false recognition that depended positively on gist memory not on verbatim memory. Gist and verbatim memory were separated with latent class analysis.
Brainerd et al. (2002)	5-adult	3	Strong vs. weak DRM lists	Age increases in false recognition and false recall that were larger for strong than for weak DRM lists.
Brainerd et al. (2004)	7–14	1	DRM and categorized lists	Age increases in false recognition for DRM and for categories. Two gist components of false memory (phantom recollection and familiarity) and one verbatim component (recollection rejection) were separated with the conjoint recognition model.
Brainerd et al. (2006)	6–14	3	DRM lists	Age increases in false recognition and false recall. Gist cuing before list presentation reduced age increases in false memory. Smaller age increases for learning disabled children than for nondisabled children. Age increases in false recognition on one-week delayed tests as well as immediate tests.
Brainerd & Reyna (2007)	6–14	2	Categorized lists	Age increases in false recognition of unrepresented exemplars of studied categories when 8 exemplars per category were studied but not when 1 exemplar per category was studied.
Brainerd et al. (2008b)	5–17	1	DRM lists	Age increases in false recognition. Gist cuing increased false memory but not true memory and reduced age increases in false memory. List repetition increased true but not false memory.
Brainerd et al. (2010)	7-adult	1	Emotional lists	Age increases in false recognition, greater false recognition for negative than for positive valence, greater age increases for negative than for positive valence. Increasing arousal amplifies the effects of negative valence.
Carneiro et al. (2009a)	3–12	2	Portuguese DRM lists with basic-level vs. superordinate critical distractors	Age increases in false recognition and false recall for basic-level critical distractors but not for superordinate critical distractors.
Carneiro & Fernandez (2010)	4–12	2	DRM lists	Age increases in false recall and false recognition when children were not warned about the illusion but age decreases when they were warned. Larger age increases in false memory with fast than with slow list presentation.
Carneiro et al. (2009b)	4-adult	3	Portuguese DRM lists with Easy vs. hard to identify themes	Age increases in false recognition and false recall. Easy to identify themes increase false memory in children and adolescents but decrease it in adults.
Carneiro et al. (2007)	3-adult	2	Portuguese DRM lists	Age increases in false recognition and false recall for lists that were separately normed for each age level. Age increases in false memory for both short and long lists.
Ceci et al. (2007)	4–9	2	Memory for common objects from a picture story	Following misinformation (suggestions that certain objects were in the story that were not), there were age increases in false memory for objects whose meanings were better understood by older than younger children.
Connolly & Price (2006)	4–7	1	Memory for the events of play sessions	Following misinformation (suggestions about play activities that children did not participate in), there were age increases in false recall for suggested activities whose gist has been repeatedly instantiated in multiple play sessions.
Dewhurst & Robinson (2004)	5–11	1	DRM lists	Age increases in semantic false recall but age decreases in phonological false recall.
Dewhurst et al. (2011)	5–11	1	Standard and phonological DRM lists and categorized lists	Age increases in false recognition for standard DRM lists categorized lists but not for phonological DRM lists. Test

Articles	Age span	# Exps. reported in article	Memory tests	Key Results
				phase priming increase false memory in older but not younger children.
Dewhurst et al. (2007)	5–11	1	DRM lists presented in story contexts vs. standard presentation	Age increases in false recognition for standard presentation but age decreases for story presentation.
Fazio & Marsh, 2008	5–7	1	Memory for false facts embedded in stories	Following misinformation (suggestions that false facts were true), there were age increases in false recall of false facts.
Fernandez-Dols et al. (2008)	6–9	3	Memory for videos and slides of people displaying positive and negative emotional expressions.	Following misinformation (suggestions that emotional had been seen that had not been seen), there were age increases in false memory for unseen emotional expressions.
Ghetti et al. (2002)	5-adult	1	Short DRM lists only.	No age changes in false recall or false recognition.
Holliday et al. (2008)	7–15	1	DRM lists	Age increases in false recall. Gist cuing increased false recall at all age levels except age 15. List repetition reduced false memory at all age levels.
Holliday et al. (2011)	7–11	1	DRM list presented as word fragments vs. standard	Ages increases in false recognition with standard presentation but age decreases with fragment presentation. False memory higher following a 3 recognition tests that after the first test.
Holliday & Weekes (2006)	8–13	1	Standard vs. phonological DRM lists	Age increases in false recognition for standard lists but age decreases for phonological lists
Howe et al. (2004)	5–12	1	DRM lists	Age increases in false recognition and false recall for maltreated, low-SES, and normal-SES children. False memory levels were not affected by maltreatment or SES.
Howe (2005)	5-adult	1	DRM lists	Age increases in false recall. Directed forgetting instructions decrease false memory in children but not in adults.
Howe (2006)	5 – 11	3	DRM and categorized lists	Age increases in false recall for word lists but not for picture lists. Equal age increases for DRM and categorized lists.
Howe (2007)	8–13	1	Standard vs. emotional DRM lists	Age increases in false recognition and false recall for standard and emotional lists. Greater age increases for emotional lists.
Howe (2008)	5–11	4	DRM lists	Age increases in false recall memory for word lists, for some lists of photographs, but not lists of line drawings.
Howe et al. (2007)	6–adult	1	DRM lists (English & French)	Age increases in false recognition and false recall in bilingual subjects. Higher false recall in children when presentation and test languages match. Higher false recall in adults when presentation and test languages mismatch. Higher false recognition at all age levels when presentation and test languages.
Howe et al. (2009)	5-adult	3	DRM lists	Age increases in false recall. Age increases are amplified when forward or backward associative is increased.
Howe & Wilkinson (2011)	7–11	1	DRM lists presented in story contexts vs. standard presentation	Age increases in false recall, with smaller increases for story presentation than for standard presentation.
Kahana & Cortese (2009)	8-adult	2	Standard vs. phonological DRM lists	Age increases in false recall visually presented standard lists. No age increases for orally presented standard lists for phonological lists.
Knott et al. (2011)	5-adult	3	DRM lists and categorized lists	Age increases in false recall. Directed forgetting instructions increased false memory in adults but not in children. Presenting list words as retrieval cues on recall tests reduced false memory.
Lampinen et al. (2006)	6-adult	2	DRM lists	Age increases in false recognition for blocked list presentation but not for random presentation. Gist cuing increased false memory in children but not in adults.
Lyons et al. (2010)	6-adult	1	Memory for causal Inferences from stories	Age increases in false memory for backward causal relations that were not presented in stories.
Metzger et al. (2008)	7-adult	3	DRM lists	Age increases in false recognition and false recall for child- and adult-normed lists

Articles	Age span	# Exps. reported in article	Memory tests	Key Results
Odegard et al. (2008)	11-adult	1	DRM lists	Age increases in false recognition when incorrect DRM list themes were cued but not when correct themes were cued. Two gist components of false memory (phantom recollection and familiarity) and one verbatim component (recollection rejection) were separated with the conjoint recognition model.
Odegard et al. (2009)	5–12	1	Memory of the events of thematic birthday parties	Following theme-consistent misinformation (suggestions that events that did not happen but were consistent with a birthday party theme had happened), there were age increases in false recognition of suggested events.
Otgaar & Smeets (2010)	8-adult	2	Dutch DRM lists	Age increases in false recall. Larger age increases when Lists were presented under a survival scenario, rather than a moving scenario.
Otgaar et al. (2012)	7-adult		Standard vs. emotional Dutch DRM lists	Age increases in false recall between 7 and 11 but not between 11 and adult. Divided attention increased adults' false memories but decreased children's. Emotional lists decreased false memory at all age levels.
Paz-Alonso et al. (2008)	8-adult	1	DRM lists	Age increases in false recognition. fMRI scans revealed age increases in false memory were associated with age changes in activation in the medial temporal lobes and the left ventrolateral prefrontal cortex.
Principe et al. (2008)	3–6	1	Memory for real-life events (a magic act)	Following misinformation (rumor mongering by another child), there were age increases in false memory for events that involved causal inferences.
Ross et al. (2006)	5–11	1	Memory for videos of a theft	Following misinformation (videos showing an innocent bystander as well as the culprit), there were age increases false eyewitness identification of the innocent bystander as the culprit.
Sloutsky & Fisher (2004)	5-adult	1	Memory for categorized photographs	False recognition of unrepresented animal exemplars of an animal increase with age when subjects were given a verbal label for the animal category.
Sloutsky & Fisher (2005)	5-adult	4	Memory for categorized photographs	False recognition of unrepresented animal exemplars of an animal increase with age when subjects were given a verbal label for the animal category.
Sugrue & Hayne (2006)	5–adult	1	DRM lists	Age increases in false recognition and false recall for standard lists but not for short lists.
Sugrue et al. (2009)	10-adult		Short vs. long DRM lists	No age trends in false recall.
Verkoeijen & Bouwmeester (in press)	8-adult		Dutch DRM lists	False recognition and false recall increased with age. Latent class lists analysis showed that level of false memory depended on level of gist processing and individual differences in gist processing were better predictors of false memory than differences in age.
Weekes et al. (2007)	9–11	1	Standard vs. phonological DRM	Lower levels of false recognition and false recall for children with low semantic processing ability on standard lists but not on phonological lists.
Wilburn & Feeney (2008)	5-adult	2	Memory for multiple photographs of the same type of animal (e.g., 10 cats)	False recognition of unrepresented animal exemplars of an animal increase with age when subjects were given a verbal label for the animal category.
Wimmer & Howe (2010)	7-adult	2	DRM lists	Age increases in false recognition. Presenting lists under divided attention conditions or deep processing instructions affected children's false memories more than adults.

Table 2

The Cornell/Cortland Emotion Lists

Valence/arousal	Property			
	Mean Valence		Mean arousal	
	List	CD	List	CD
Negative/high				
<i>Cold</i>	5.30	4.02	4.50	5.19
<i>Hurt</i>	2.96	1.9	5.72	5.85
<i>Thief</i>	3.61	2.13	5.66	6.89
<i>Spider</i>	4.64	3.33	5.16	5.71
<i>Dead</i>	4.18	1.94	4.94	5.73
<i>Rough</i>	5.29	4.74	4.16	5.33
<i>Anger</i>	3.73	2.85	7.10	7.17
<i>Needle</i>	4.48	3.82	4.88	5.36
Grand means	4.27	3.09	5.27	5.90
Positive/High				
<i>God</i>	6.25	8.15	5.58	5.95
<i>Baby</i>	6.14	8.22	5.06	5.53
<i>Love</i>	7.10	8.72	4.82	6.44
<i>Music</i>	6.14	8.13	4.54	5.32
<i>City</i>	5.40	6.03	4.98	5.24
<i>Beach</i>	6.78	8.03	5.13	5.53
<i>Hug</i>	6.94	8.00	4.73	5.53
<i>Pretty</i>	6.29	7.75	4.99	6.03
Grand means	6.38	7.88	4.98	5.70
Negative/low:				
<i>Sick</i>	3.91	1.9	5.04	4.29
<i>Fat</i>	5.54	2.28	4.55	4.81
<i>Shy</i>	5.00	4.64	4.98	3.77
<i>Dirt</i>	4.85	4.17	4.39	3.76
<i>Board</i>	5.75	4.82	4.99	3.36
<i>Trash</i>	4.15	2.67	4.43	4.16
<i>Glass</i>	5.73	4.75	4.66	4.27
<i>Fall</i>	4.59	4.31	4.94	4.70
Grand means	4.94	3.69	4.75	4.14
Postive/low				
<i>Soft</i>	5.77	7.12	4.40	4.63
<i>Green</i>	5.28	6.18	4.26	4.28
<i>Nice</i>	6.36	6.55	5.19	4.38
<i>Sleep</i>	6.12	7.20	3.70	2.80
<i>Bird</i>	6.51	7.27	4.47	3.17
<i>Grass</i>	6.31	6.12	4.57	4.14

Valence/arousal	Property			
	Mean Valence		Mean arousal	
	List	CD	List	CD
<i>Fish</i>	6.07	6.04	5.30	4.00
<i>River</i>	6.25	6.85	4.42	4.51
Grand means	6.08	6.67	4.54	4.00

Note. List = the 15 words on each list, and CD = the critical distractor for each list. For the 15 list words, some of the words on each list are not available in existing word norms for valence and arousal (Bradley & Lang, 1999; Toggia & Battig, 1978). The tabled means for valence and arousal of list words are therefore based on the words that are available in these norms.