Representational Constraints on the Development of Memory and Metamemory: A Developmental–Representational Theory

Stephen J. Ceci
Cornell University

Wendy M. Williams
Cornell University

Stanka A. Fitneva
Queen’s University

Traditional accounts of memory development suggest that maturation of prefrontal cortex (PFC) enables efficient metamemory, which enhances memory. An alternative theory is described, in which changes in early memory and metamemory are mediated by representational changes, independent of PFC maturation. In a pilot study and Experiment 1, younger children failed to recognize previously presented pictures, yet the children could identify the context in which they occurred, suggesting these failures resulted from inefficient metamemory. Older children seldom exhibited such failure. Experiment 2 established that this was not due to retrieval-time recoding. Experiment 3 suggested that young children’s representation of a picture’s attributes explained their metamemory failure. Experiment 4 demonstrated that metamemory is age-invariant when representational quality is controlled: When stimuli were equivalently represented, age differences in memory and metamemory declined. These findings do not support the traditional view that as children develop, neural maturation permits more efficient monitoring, which leads to improved memory. These findings support a theory based on developmental–representational synthesis, in which constraints on metamemory are independent of neurological development; representational features drive early memory to a greater extent than previously acknowledged, suggesting that neural maturation has been overimputed as a source of early metamemory and memory failure.

Keywords: memory development, metamemory, PFC, binding, representation

The traditional account describes memory development as a confluence of the growth of four factors: strategies, knowledge, metamemory, and capacity (processing speed). Particularly over the early and middle childhood years, there is dramatic growth in these factors. Although strategies can be successfully taught and have been shown to increase memory performance (Ornstein, 1978) and knowledge can be taught and likewise has been shown to enhance performance (Ceci, 1980; Chi & Ceci, 1987; Ornstein & Naus, 1985), metamemory and capacity are assumed to be under strict maturational control and, hence, not readily teachable. These same four factors have been part of the canonical developmental account for nearly forty years, as Ornstein, Haden, and Hedrick (2004, p. 374) remarked,

At the 1971 meeting of the Society for Research in Child Development, John Flavell organized a now-celebrated symposium on children’s memory. The title of that symposium—“What is memory development the development of?”—posed a question that as we see it, has defined the nature of research on children’s memory for more than 30 years. Changes notwithstanding, there has been remarkable constancy in that the major thrust of research has most often been on the characterization of something (e.g., memory capacity, mnemonic strategies) that is undergoing development.

The developmental–representational theoretical synthesis preferred here extends this traditional account in several ways. Most importantly, it demonstrates that knowledge and metamemory function in parallel, with advances in the former enhancing the efficiency of the latter. Ordinarily, young children’s knowledge structures contain less information, and what information they do contain is less organized; however, there are domains in which young children’s knowledge is superior to older children’s, and in such domains, metamemorial efficiency improves. Thus, even if metamemory is under maturational control, it can be precociously ratcheted up by the availability of rich knowledge structures, and likewise it can be degraded in older individuals by the presence of nonoptimal knowledge structures.

Strategies and processing speed will normally work against young children (Kail & Ferrer, 2007), and therefore, any gains that accrue from the availability of richly structured knowledge will to some extent be offset by age differences in strategy deployment.

In his critique of memory development research, Siegler (2004) argued that there was insufficient attention given to the underlying
mechanisms that led to age differences in mental processing: “Memory is memory of mental activity. Events in the external environment cannot be remembered; all that can be remembered is our processing of them” (p. 474). The developmental–representational synthesis described in this article provides a framework for thinking about developmental differences in processing. It posits that during encoding, the child’s informational environment becomes assimilated into existing knowledge structures. This, in turn, enables the child to notice similarities and differences with extant knowledge and to infer shared attributes across items to form categories (e.g., citrus, predator, dairy), both of which influence the child’s skill at mentally patrolling these encodings.

Hence, the developmental–representational theory shares with traditional accounts a recognition of the role of the same four factors in remembering, but it reconstitutes the manner in which they affect memory by providing evidence that they catalyze each other. It does not deny that capacity differences exert limitations on very young preschoolers (see Kail, 2008) or that on average, older children possess more knowledge and strategies. By virtue of its focus on the nature of the mental representation, it leads to targeted expectations that traditional accounts do not, such as reverse developmental effects and precocious metacognition. In so doing, it calls into question a longstanding set of assumptions about the role of neural immaturity in metacognitive growth, especially its role in the monitoring of episodic memory. We contrast this synthesis with current theories of memory development (trace integrity, fuzzy trace, binding theory), with which it shares some but not all assumptions (e.g., more integrated traces with age).

Introduction

Traditional developmental accounts of memory growth during early and middle childhood entail a set of linear assumptions: As preschool children grow older, the brain structures responsible for monitoring and regulating memory and metamemory mature; this brain maturation underpins improved monitoring and regulatory ability, which in turn, enhances memory accuracy. According to this set of assumptions, neurological influence on metamemory constitutes a formidable constraint on memory growth: If the brain regions responsible for effective monitoring and tracking of encodings are undeveloped, both children’s awareness of their recall readiness and the strategies needed for mnemonic intervention will be lacking, thus leading to poor memory performance. There is an intuitive appeal to this set of assumptions, and empirical data exist to support each assumption in explaining the inferiority of very young children’s recognition and recall performance.

This article poses a theoretical synthesis not previously proposed or tested. Challenging the role of neurological constraints on early metamemory and memory development, we provide argument and data suggesting that the nature of memory representations are the primary driving force in memory development in early childhood. We argue that neurological development, although clearly foundational and important for a wide variety of cognitive attainments, is already at a sufficient level during early childhood and, thus, is itself insufficient to explain much of the age-related variance in metamemory and recognition accuracy exhibited during this period of development.

Four Characteristics of the Proposed Developmental–Representational Theory

First, we propose a conceptual framework that posits a mechanism through which developmental variation in both recognition memory and metamemory performance comes about. This mechanism consists of the interaction of mental representations with metacognitive insights.

Second, we argue that the ability to effectively deploy metamemory (the awareness, monitoring, and regulation of the contents of memory) is a function of the nature of the memory representation. If correct then a strong form of the argument is that mental representations play a fundamental mediational role, such that when representational quality is equated across ages, metamemory efficiency and recognition memory accuracy will not vary by age.

Third, we describe an important corollary, which is that the role of neurological maturation is overestimated in theories that fail to take into consideration the possibility that representational quality mediates memory and metamemory development. Note that this is not the same as claiming neurological changes are unimportant in recognition and metamemory development. Rather, our claim is that age-related variation on memory and metamemory tasks often occurs independent of the influence of neurological changes and, therefore, that such variation can be shown to occur when neurological status is held constant. A somewhat weaker argument following from this corollary is that the level of neurological development needed for effective memory monitoring is achieved early on, and thus, neurological changes that happen in the preschool and early elementary school years do not underlie children’s metamemory capacity.

Fourth, we note that the proposed synthesis focuses on a thorny issue in memory development, namely, situations in which canonical developmental trends are either absent or reversed (Brainerd, Reyna, & Ceci, 2008). Such reversals, albeit rare, present a serious challenge to theory, and past accounts of their occurrence have been post hoc and have not led to a priori predictions of when younger and older children’s performance will be similar or reversed. We offer a strategy for predicting when and where younger children’s memory and metamemory functioning will be inferior to, equivalent to, or even superior to older children’s and adults’.

Ultimately, targeted expectations will benefit from development of measures of mental representations that can be integrated across levels, from neurological (e.g., unique patterns of hippocampal cell firing) to cognitive (e.g., semantic proximity) and behavioral.

In this series of experiments, we present data consistent with such a theoretical synthesis by holding chronological age—a proxy for normative neurological maturation—constant, while allowing representational factors to vary both within and between children. To adumbrate our conclusions, the proposed developmental–representational theory receives considerable support. Results indicate that when children of different ages have equivalent mental representations for some items but different representations for others, they exhibit equivalent metamemory (and more similar recognition memory) for the former, while exhibiting traditional age-related trends for the latter. Without something like the pro-
posed theory, it would be unclear why younger children appear to lack the ability to engage in metamemory on some trials but not on others that are nominally identical. And it would be unclear why age differences in recognition disappear in some situations while remaining in evidence in others that are highly similar, given the equivalence of neurological status across situations.

Background

By the early 1990’s, a consensus had emerged about basic developmental trends in remembering and forgetting. Brainerd, Howe, and their colleagues (Brainerd & Reyna, 1990; Brainerd, Reyna, Howe, & Kingma, 1990; Howe & Brainerd, 1989; Howe, Brainerd, & Kingma, 1985), as well as Bender, Wallsten, and Ornstein (1996) and others, reported discoveries about the shape of the developmental forgetting and retention curves and, more important, the theoretical processes that underpin them (for reviews see Kail, 1996; Schneider & Pressley, 1997). In addition to age-related changes in how much information children remember, there was a growing awareness of the loci of such changes (i.e., whether they were rooted in processes undertaken during encoding, storage, and/or retrieval) as well as the decay functions associated with loss of information at each locus (e.g., Brainerd et al., 1990; Howe et al., 1985; Howe & Brainerd, 1989).

A set of three observations animated much of this research: (a) sometime between middle childhood and adolescence, children acquire adultlike metamemory, that is, the skills of monitoring, introspecting, and patrolling the contents of their memory; (b) there are substantial neurological changes in the frontal cortex and especially in the prefrontal inferior region and the medial temporal area that subserve important cognitive developments (e.g., Rolls, Stringer, & Trappenberg, 2002; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005), including metamemory, and these neurological changes accelerate between the ages of 3 years and 6 years; and (c) coincident with these rapid developments in both the metamemory and the frontal cortices are increases in recognition memory performance tied to processes undertaken at encoding and retrieval. A plausible assumption was that these three developments stand in a causal relationship, with neurological changes in the prefrontal region enabling more efficient monitoring and regulating of the contents of memory. This, in turn, leads to better memory performance because it brings to awareness mental interventions needed to enhance recall and recognition readiness. We refer to this as the traditional developmental account of memory and below review support for each of these three planks and their linkages.

Links Between the Development of Metamemory and Memory Changes

An aspect of memory development that emerged as an important explanatory construct during the past quarter century is metamemory—the knowledge about how one’s memory functions and the ability to introspect into the workings of one’s memory, to monitor and regulate its effectiveness. Metamnemonic knowledge refers to insights about tasks, strategies, or internal cognitive abilities. Metamemory regulation refers to two processes involved in the coordination of memory: monitoring and control. Monitoring entails the introspective awareness of an item in memory, and control entails the manipulation of this item to maximize performance (Nelson & Narens, 1990).

Researchers have provided a great deal of descriptive data about the development of various metamemory subprocesses involved in monitoring and regulating memory and their causal role in recall and recognition performance (for seminal demonstrations, see Flavell & Wellman, 1977; Wellman, 1978). In an extensive review, Schneider (1999) described developmental trends in the use of various mnemonic strategies and metamemory knowledge, showing that preschool-aged children are notably deficient in the use of most. There is some experimental evidence showing that children as young as 4 years have insights into both the contents and the workings of their memory (Balcom & Gerken, 2008; Caltice, Somerville, & Wellman, 1983), but such insights are rudimentary and continue to develop throughout late childhood and into adolescence (e.g., Zabrocky & Ratner, 1986). Although 5-year-olds to 6-year-olds possess more metamemory knowledge than do younger children, it is not until adolescence that such awareness reaches adultlike levels (Flavell, 1999).

Furthermore, there is a moderate size relationship between metamemory ability and children’s actual memory performance (r = .41 across 60 studies analyzed by Schneider & Pressley, 1997). Even among very young preschoolers, those who monitor better remember more (Balcom & Gerken, 2008). Intuitively, this suggests that once a child is capable of monitoring the contents of memory, it is possible to intervene to shore up weak encodings (e.g., by mentally rehearsing), assess recall-readiness (by self-testing), appreciate risky circumstances (e.g., ones that lead to buildup of proactive inhibition), and so on. Absent such monitoring and metamnemonic awareness, the child may not realize that the to-be-remembered material has not yet been committed to permanent storage, is currently inaccessible, or is vulnerable to interference. Thus, metamemory development and memory performance appear to be not only correlated but causally linked, and empirical evidence supports this contention.

Koriat, Goldsmith, Schneider, and Nakash-Dura (2001) examined age differences under various recognition and recall formats that were designed to differentially require children to monitor the contents of their memory. They demonstrated that children’s memory performance “depends critically on the ability to monitor the veracity of the information that comes to mind and to regulate memory reporting in accordance with both the monitoring output and the operative incentives for accuracy” (p. 435). In short, young children’s presumed deficits at monitoring the contents and operations of their memory is often explicitly assumed to be a cause of their inferior memory performance vis-à-vis older children, who possess greater metamemory knowledge and engage in more effective monitoring.

Links Among the Development of Metamemory, Memory, and Brain Development

The immaturity of frontal neural systems is assumed to constrain metamemory and thus to be a key cause of age-related differences in memory (e.g., Schacter, Savage, Alpert, Rauch, & Albert, 1996). As will be seen, there is ample evidence demonstrating that frontal lobe damage is specifically associated with impairment of metamemory (Shimamura, 1985), and a number of studies have assessed metamemory monitoring in frontal-lobe-
lesioned patients and shown these patients to be deficient (Janowsky, Shimamura, & Squire, 1989). Several functional magnetic resonance imaging studies have reported activations in frontal regions associated with metamemory behaviors, such as feeling-of-knowing judgments (Kikyo, Ohki, & Miyashita, 2002; Maril, Simons, Mitchell, Schwartz, & Schacter, 2003). As Souchay and Isingrini (2004) note, there is a great deal of evidence linking metamemory to “regions throughout the prefrontal cortex (bilateral inferior frontal gyri, left middle frontal gyrus) and anterior cingulate cortex. Thus, evidence from neuropsychological and imaging studies suggests that the frontal cortex may be important for metamemory monitoring” (p. 90).

In this section, we briefly review the evidence linking frontal maturation to memory via the mediation of metamemory. The aspect of metamemory that has received the most attention by developmental scientists concerns age-related differences in the awareness that one does not remember. Developmental studies of young children’s memory monitoring (e.g., Flavell, Friedrichs, & Hoyt, 1970), animal studies of tracking past memories (Hasegawa, Blitz, Geller, & Goldberg, 2000), and research with elderly patients suffering from pathological conditions (Bauer, Kyaw, & Kelbey, 1984; Schacter, 1996; Souchay & Isingrini, 2004) all show a lack of awareness about what is no longer remembered. In these experiments, participants are unaware that they cannot recollect some item or event, and as a result, they volunteer an answer that is incorrect, rather than recognizing that they do not remember. Developmental work on this phenomenon emanates from the classic studies of Flavell and his colleagues. For example, when asked to estimate how many words they can recall from a list or to raise their hands when they have mastered an entire list of words, young children are typically poor at assessing the current status of their memories: They claim they can recall words that they have not yet committed to memory (Flavell et al., 1970). In comparison, because older children are better able to monitor the contents of their memories, they know with greater exactitude when they do not yet remember. This enables them to intervene with appropriate strategies (see Schneider, 1999).

Developmental research on young children’s lack of awareness about the contents of their memories finds a parallel in cognitive neuropsychological research in which the lack of awareness that something cannot be recalled implicates the frontal cortices. Numerous studies have reported correlations between PFC activity and memory monitoring efficacy. In an early and influential review of this correlation, Tulving and his colleagues (Wheeler, Stutt, & Tulving, 1997) summarized lesion studies as well as positron emission tomography and functional magnetic resonance imaging studies of healthy participants, noting the remarkable consistency across findings showing that PFC regulates episodic memory functioning by establishing, maintaining, and monitoring neurocognitive programs. A number of researchers have documented that right dorsolateral PFC is among the PFC regions that supports processes engaged during monitoring and/or evaluating the products of retrieval attempts (e.g., Cruse & Wilding, 2009; Hayama & Rugg, 2009; Henson, Shallice, & Dolan, 1999).

Both animal and human studies have demonstrated that a tonic signal in the PFC changes over time, allowing the healthy organism to predict and monitor changes in memory (Hasegawa et al., 2000). In addition, clinical studies with patients who are unaware that they have forgotten the information requested of them reveal the involvement of the PFC in their forgetting. In discussing patients with global retrograde amnesia (temporary forgetting), Schacter (1996, p. 92) remarked that they have difficulty “remembering they forgot.” For instance, when asked what they ate for breakfast, such patients routinely forget that they cannot remember and often claim they ate things they did not (e.g., Schacter, Curran, Gallucio, Milberg, & Bates, 1996). They do not monitor their memory to realize that they ate at breakfast is (at least temporarily) inaccessible. Bauer et al. (1984) reported Korsakoff patients with frontal-lobe problems had less accurate metacognitive control due to a frontal lobe deficit. A positron emission tomography study with normal participants demonstrated that tasks requiring the most left PFC activity involved the greatest degree of metacognition, whereas lists that were already organized were associated with the least activity in that region, suggesting that metamemory functions are mediated by the PFC neural system (Fletcher, Shallice, & Dolan, 1998). In a related case, on the basis of imaging data, Fernandez-Duque, Baird, and Posner (2000) and Shimamura (2000) have also suggested that metacognition is closely related to executive functions supported by the PFC.

At the other end of the developmental spectrum, Souchay and Isingrini (2004) have demonstrated that age-related declines in metamemory control result from executive limitations associated with deterioration of the PFC among elderly. Using event-related potentials, numerous investigators have established the prefrontal-dependence of functions that underpin metamemory, especially monitoring, such as selective attention and recency judgments (see Hackman & Farah, 2009; Stevens, Lauinger, & Neville, 2009).

Given the evidence reviewed above for the involvement of PFC structures in the growth of memory and metamemory, early preschool-aged children’s difficulty in monitoring the contents of their memory has been assumed to be the result of immature brain development. The maturation of the PFC extends over a long period of development, from initial synaptic proliferation during infancy followed by a pruning until maturation during early adolescence, with nonlinear changes in gray matter density through the postpubertal period (Giedd et al., 1999; Sowell, Delis, Stiles, & Jernigan, 2001). Nelson and his colleagues (Thomas & Nelson, 2001) have shown that the physiological maturation of widespread neural networks in the PFC that are correlated with executive functions and working memory emerge over protracted period of development, with 8-year-olds’ PFC development superior to 4-year-olds (Luciana & Nelson, 1998).

It has been shown that the maturation of the prefrontal inferior region (along with the medial temporal area) is not only associated with metamemory monitoring but is also important for success on memory tasks tapping these abilities (see Schacter, Normal, & Koutstaal, 1998, for review). As Luria (1973) documented long ago, children between the ages of 4 years and 7 years undergo rapid changes in the frontal areas, thus suggesting that the youngest children may be likely candidates for any type of forgetting that is subserved by these regions (Welsh, Pennington, & Gruesser, 1991). Furthermore, neuroimaging studies have consistently shown that the PFC plays a crucial role in establishing the decision criteria used by participants to judge whether a subset of features belongs to a specific episodic memory (Johnson, Kounios, & Nolde, 1996; Wilding & Rugg, 1996). Young children have been shown to use different decision criteria in deciding whether a previously presented stimulus is familiar, thus achieving accuracy.
at the expense of quantity, of how much they remember (Koriat & Goldsmith, 1996).

In sum, it is well-established that PFC-mediated memory is developmentally sensitive, and damage to the PFC can disrupt thereassembling and monitoring of memories and memory functions more generally (Janowsky et al., 1989). There are strong reasons to believe that the prolonged development of the PFC strongly influences metamemory mechanisms and, ultimately, influences memory performance itself. Figure 1 illustrates the causal role assigned to this brain region in explaining developmental differences in processing efficiency, in this case monitoring. In the figure, there are steep developmental gradients in monitoring efficiency for all kinds of representations, from ones that young children know a great deal about (e.g., Sesame Street characters) to ones they know very little about (weather). Thus, in the traditional view, immature brain development constrains cognitive processing of all events, regardless of their nature, and it is the processing inefficiency (again, in this case monitoring) that underlies poor memory performance of young children and brain-damaged adults. In the figure, processing of information that is familiar to children is shown as somewhat more efficient than less familiar information, given the well-established advantages of encoding and retrieving familiar information (e.g., Chi & Ceci, 1987). However, the monitoring advantages associated with familiarity are dwarfed by those associated with neural maturation.

In the following pilot study and four experiments we probe the phenomenon of young children’s failure to recognize items whose contextual attributes they are able to recognize. These experiments examine an alternative set of hypotheses about the sources of memory and metamemory development in early and middle childhood rather than the traditional PFC-mediated account depicted in Figure 1. As a consequence of this exploration, we arrived at a new synthesis, called the developmental-representational theory, which is described in the General Discussion.

Pilot study. In contrast to those who volunteer an answer that is incorrect rather than recognizing that they do not remember (Schacter, Curran, et al., 1996), the present research focuses on the flip side of the problem—namely, not recognizing that remembering has occurred. Are there age-related differences in children’s understanding and awareness of the contents of their memory when they may in fact have a memory that is potentially accessible?

A pilot study sensitized us to this possibility. We presented preschool-aged children with a series of pictures of common objects (e.g., animals, toys, transports) and told them they would be asked later whether they remembered the objects. We gave them two metamemory tasks, one at the time of the presentation of these pictures and the other 2 days later at the time of the subsequent recognition task. The first metamemory task was an ease-of-learning task in which children were asked which of a pair of presented pictures would be easiest to recognize later. Ease-of-learning judgments are better predictors of adults’ subsequent recognition performance than other types of metamemory (Leone-sio & Nelson, 1990), although such tasks are difficult for young children; the second metamemory task was a remember–know judgment at the time of later recognition testing. When we asked children to recognize the pictures interspersed among foils, they often claimed they could not. Yet, they were subsequently able to correctly guess the context in which the pictures had been located during their initial presentation (e.g., the color of the background poster on which the picture had appeared or the spatial location in the house drawing in which it appeared). Barring unusual circumstances, adults would reason that a picture must be old if its contextual attributes seem familiar. It seemed these young children appeared to remember the context of the very items that they claimed not to remember having seen. (Later, we review neurological models that shed light on such dissociations.)

Thus, the classic source memory failure—in which children recognize an event as familiar but fail to correctly identify its attributes, such as its source (e.g., Lindsay, Johnson, & Kwon, 1991), is the opposite of what was observed in this pilot study; that is, children identified the picture’s attributes without recognizing the picture itself. It is interesting to note that in contrast to previous work with adults, neither ease-of-learning judgments nor remember-know judgments that were adapted for young children predicted recognition accuracy.

Encoding specificity? Three decades ago, Tulving and his associates (Muter, 1978; Tulving & Thomson, 1973; Watkins & Tulving, 1975; Tulving & Wiseman, 1975; Wiseman & Tulving, 1976) reported a series of experiments in which college-aged

---

**Figure 1.** Prefrontal cortex (PFC)-driven model of episodic memory monitoring. Regardless of how rich a child’s representation is, an adolescent will always monitor that item more effectively because of their greater neural development. Stills from Sesame Street reprinted with permission. Sesame Workshop and Sesame Street, which are registered trademarks, and associated characters, trademarks, and design elements are owned and licensed by Sesame Workshop. Copyright 2010 by Sesame Workshop. All rights reserved. Photo credit: John Barrett.
participants recalled words that they had previously failed to recognize. For example, participants in Tulving and Thomson’s (1973) study encoded the word chair in the context of the word glue (glue–chair). After studying the word pairs, many participants did not recognize the target word chair when they saw it presented alone, but they were able to recall it when given the cue “glue,” a finding that was interpreted in terms of encoding specificity. Although this is an episodic task, Tulving’s colleagues demonstrated a similar failure to recognize recallable words with semantic tasks. For example, in one study participants saw the word Bryan in a list of names. They might encode Bryan as a first name or a surname for a number of individuals known to them. Many of the participants in these experiments failed to recognize the word Bryan when it was presented as part of a standard recognition test, but later they were able to recode it when presented with the recall cue “William Jennings ____” (Muter, 1978). Again, this was interpreted as the result of encoding specificity creating a mismatch between two functionally different traces: participants failed to retrieve the identical nominal word because the encoded word differed from it. However, these examples of episodic and semantic failures to recognize recallable items bear little relation to the phenomenon observed in the pilot study. In the latter, children failed to recognize the same episodic item whose spatial cues led them to retrieve (e.g., failing to recognize that a bicycle had been presented but when asked to guess in which room it had been located or what color the background posterboard was, they correctly recalled these attributes). This is a deeper dissociation than that produced in the encoding specificity experiments because it suggests that the child is able to access attributes of the very same encoded trace they failed to recognize, not simply a different encoded trace of the same nominal stimulus. Of course, this hinges on replicating the pilot results and ruling out alternative explanations such as recoding of the original trace at the time of cueing with its attributes.

We hypothesize that what develops is not the ability to introspect into the contents of one’s memory—that is, metamemory—but rather the ability to represent the contents of memory in a manner that makes it accessible to introspection and monitoring and, hence, successful cueing. We further hypothesize that metamemory is potentially available long before it gets deployed consistently, because of representational constraints on its operativity. Inaccessible memories highlight the potential role of representational constraints because such memories are by definition available but not readily retrievable, possibly due to how they are represented (e.g., fewer or less well-integrated attributes that fail to activate each other).

The above hypotheses entail disconnecting metamemory from the nature of the mental representation and presents the latter as an alternative causal variable, one that mediates the efficacy of the former. Notwithstanding studies showing young children’s inefficient metamemory linked to their poor memory performance (Schneider & Pressley, 1997) and notwithstanding the cognitive neuroscience studies showing a causal link between various regions of the PFC and metamemory performance, young children’s poor recognition memory may turn out to be critically related to the way information is represented in memory, which in turn affects the ability to engage in effective metamemory processes such as monitoring. Thus, preschoolers may have the ability to introspect and monitor the contents of their memories if their memories are represented richly enough to allow access to metamemory. Conversely, young children may fail to recognize an item that is poorly represented in memory, even if they are able to recognize its contextual attributes. According to such an account, brain maturity may be a necessary but not sufficient condition for effective metamemory functioning to occur, and it may come online quite early in development, so that it accounts for little variance in these simple recognition tasks when the nature of the mental representations on which metamemorial processes operate are controlled.

This leads to a set of interrelated hypotheses: In early childhood, the mental representations for many items are impoverished in terms of their number of attributes, their degree of integration, and the Euclidean distance between an encoded item and related items—and it is this representational impoverishment that constrains the efficacy of metamemory, rather than brain immaturity per se, which although necessary is not sufficient and may come online much earlier. When the mental representations of young children are not impoverished (e.g., when they have similar numbers of attributes as do older children’s representations), children’s ability to introspect into the contents of their stored traces should be sufficient to judge an item as familiar on the basis of its contextual attributes, even if the item-level activation is insufficient. (Later, we test this hypothesis by manipulating item representations of younger and older children and assessing their failure to recognize items whose contextual attributes were recognized as well as their metamemory efficacy. Thus, we test this hypothesis not by correlating neural activation during encoding with subsequent memory performance but rather by holding it constant while varying representational complexity.)

Some empirical findings hint at the viability of a representational account of early memory development. For example, on memory tests, young children perform better on recognition tests than recall tests, due in part to the representation of the original retrieval cues, that is, the picture itself (Holliday, Brainerd, & Reyna, 2008). Kobasigawa (1974) reported evidence of an age-related increase in the efficient employment of available retrieval cues (i.e., cue cards depicting contextual attributes) from 5 years to 11 years of age. The younger children encoded as many contextual attributes as did the older children; however, they did not spontaneously use this available contextual information to assist with recall. Recently, Ceci, Papierno, and Kulkofsky (2007) reported a significant relation between the Euclidean distance between an encoded item and its distractor and children’s performance on a recognition task. They analyzed incidences of failing to recognize original items coupled with falsely recognizing these items’ distractors and found that as the similarity between items and their distractors decreased, false recognition of the distractor decreased. This was due to the far greater negative effect of a semantically proximal distractor on recognition accuracy than a more semantically distal distractor. Because the effect for similarity did not interact with age, this suggests that when knowledge is represented similarly for younger and older children, the former are no more likely to confuse distractors for original items than are older children.

It might seem odd for children to be able to recognize an item’s attributes but not recognize the item itself. However, recent neuroscience research strongly suggests that different encoding processes that emerge in distinct brain regions are involved in pro-
cessing item familiarity and contextual attributes and that an item’s contextual attributes can be recognized even when the item itself may not be. Specifically, hippocampal activation during item encoding does not predict later recognition of the item, even though it does predict later recognition of its contextual attributes (real vs. imagined; male voice vs. female voice). On the other hand, perirhinal activation during item encoding predicts later recognition of the actual item itself but not the recognition of its contextual attributes (Davachi, Mitchell, & Wagner, 2003; Davachi & Dobbins, 2008). Several investigators have reported that adults show source memory for items they do not recognize and for items they claim elicit no memory for source detail (Starns, Hicks, Brown, & Martin, 2008; Wais, Mickes, & Wixted, 2008). For adults this occurs because of a conservative old–new decision criterion (which is associated with a very low false alarm rate). Whereas for young children in our pilot study (and for the subsequent studies to be reported as well), this phenomenon occurs even though they have a very high false alarm rate. Thus, the adult findings are a different phenomenon from what will be reported here, and something other than a conservative decision criterion is at work for young children.

Finally, Barsalou (2008) noted an emerging consensus in neuroscience that categorical knowledge resides within encapsulated modules (e.g., visual, auditory). When a child perceives a picture of an animal, for instance, its full representation does not reside in a transdomainal, amodal semantic memory but rather is distributed across domains that may be only partially activated by the provision of only one of its attributes.

What these findings by Davachi et al. (2003) and Barsalou (2008) have in common is that they lead to the expectation that items may fail to be recognized as a result of attributes stored in medial temporal lobe structures that are not brought into awareness during familiarity judgments controlled by the PFC. One of the goals of the present study is to ask whether this is true in general or only when the representation is too impoverished to support metamemorial awareness of the contents of memory.

Synopsis of the Present Research

Experiment 1 refined and extended our pilot study. When queried about previously experienced events, we inquired whether young children would claim they have no memory of such events, even though they can recognize their contextual attributes. Given evidence of failing to recognize items whose attributes have been recognized in this first experiment, Experiment 2 was designed to rule out a procedural artifact that could have produced the findings in both Experiment 1 and the pilot study. In Experiments 3 and 4, we examined the hypothesized mechanism underlying this form of recognition failure; specifically the interdependence of the ability to recognize events, the ability to monitor memories of such events, and the representation of such events in memory. The logic of these latter experiments is straightforward: We provided controls for age differences in the type and amount of information that children had available, to determine whether such differences were at the root of developmental differences in both recognition memory and metamemory. By controlling for representational differences, we can ask and answer the question that lies at the heart of this program of research: Is younger children’s poorer recognition memory the result of inherently poor metamemory, possibly the result of immature frontal lobe development? Alternatively, is metamemory itself mediated by the structure and richness of the representation, independent of neurological development factors? Again, the latter claim does not imply neurological changes are unimportant in the course of memory development, but it does imply that they may have become sufficient earlier in development and that representational factors constrain memory long after neurological conditions are sufficient for monitoring and recognition. If correct, this view would lead to overimputations of the role of brain mechanisms in theories about children’s emerging memory and metamemory.

To summarize, the goal in the present research was to determine whether recognition failures in conjunction with readily available episodic information are due to pervasive metamemory deficits (i.e., a neurologically based limitation on the ability to monitor the contents of memories) or whether such failures result from representational constraints that are independent of brain developments. In addressing this possibility, we examine the hypothesis that younger children’s recognition memory and metamemory may be on par with older children’s, if and when the items to be remembered are equivalently represented.

**Experiment 1**

Experiment 1 replicates and builds on our pilot findings. The procedure here and in the following two experiments was similar to that of the pilot study. We presented younger and older preschool children—those known to be below and above the onset of early metacognitive awareness—with a series of pictures of familiar toys, animals, and transports. These items appeared in a context: (a) a posterboard of a given color, (b) with four houses, one in each quadrant of the posterboard, (c) each containing four rooms. Children’s recognition memory for the pictures was assessed (“Is this one of the pictures you saw?”), as was their memory for each picture’s three context attributes: room location, poster quadrant, and posterboard color. These three levels were probed by asking (a) in which one of the four rooms in this house did the picture appear (see Figure 2), (b) in which one of the four quadrants on the posterboard did this house appear, and (c) what color was the posterboard on which it appeared? The purpose of this repeated probing was to determine whether memories for the

![Figure 2](image-url) Illustration of the stimuli used in the experiments. Four such houses, with different items, were shown on each of the four differently colored poster boards. There was one house in each of the quadrant of a board.
context of pictures children had previously claimed not to recognize differed in both type and amount of information compared with memories for pictures they did recognize.

On the basis of our pilot work, we expected that previously recognized pictures would be more likely to be associated with their contextual attributes than would pictures that were unrecognized. Additionally, we were interested in whether particular levels of contextual information were differentially associated with accurate recognition. Specifically, the most proximal location—room in the house—may be more closely associated with picture recognition than are more distal ones (color of posterboard background). This is because, phenomenologically, it would seem harder to remember the immediate setting of an item and yet not remember the item itself. As Starns and Hicks (2008) have shown, contextual information of the form used here (color and location attributes) are bound to the item more powerfully than they are bound to each other. Hence, we wondered whether children sometimes retrieve above-chance amounts of information about pictures they claimed not to have seen previously and whether this would occur more so in response to distal cues than proximal ones?

**Method**

**Participants.** Fifty-four preschoolers participated in the first experiment, 27 younger ones ($M = 43$ months, $SD = 3.2$ months) and 27 older ones ($M = 70$ months, $SD = 5.1$ months). Fifteen of the younger children and 14 of the older ones were girls. Children were recruited from regional day care centers, nursery schools, and kindergartens that enrolled children from working class and middle class families. Three of the younger children’s families were receiving free or reduced lunch temporary assistance to needy families payments (welfare, TANF) at the time of the study, as were two of the older children’s families.

**Procedure.** The data were collected over 2 consecutive days. The first day was devoted to presenting the pictures and the second day was devoted to testing recognition accuracy and recognition of contextual attributes.

On the first day, children were introduced to four large, differently colored posterboards, one at a time. On each of these posterboards, there were four 4-room houses, one in each of the posterboard’s four quadrants. A picture of a common toy, animal, or transport was located in each of the four rooms of each of the four houses (see Figure 2). Thus, there were 64 pictures in all (4 posterboards $\times$ 4 houses on each posterboard $\times$ 4 rooms in each house). Children were given two practice trials with houses located on white and black posterboards that were not used in the main experiment. They were told:

> I want to show you drawings of some houses. In each house there are four rooms (showing a poster and indicating each of the four houses and the rooms in each), and there is a different picture in each of these rooms that I want you to look at, OK? This is fun, so let’s look at them together. (First pointing to the house in one of the quadrants of a white posterboard) Let’s see what pictures are in this house, OK? Upstairs I see a picture of a slinky in one room. And a picture of a wagon is in another room. Now let’s look downstairs. I see a room with a picture of a toy elephant in this room, and a picture of a kite in this room. Do you see them, too? Good. Now let’s look at another house on this poster sheet and see what’s in its rooms, OK? Let’s see if you can name each thing, all right?

After the pictures in the four rooms of a given house were named by the child, that particular house was removed from view. This was done so that presentation time for the pictures in each house was roughly the same, preventing children from dwelling on some houses at the expense of others. Following the first house, the next house on that posterboard was shown, starting at a different quadrant, and each of its toy pictures was named before being removed from sight, and so on, until all 16 of the pictures in each of the four houses on that posterboard were named. At this time, the posterboard was removed from view and another was introduced. This procedure was repeated until all 64 pictures were shown and named. The location of pictures on the posterboards was randomized across children, and the starting quadrant was rotated across children. Effort was taken to ensure that no picture bore an association with a particular room, for example, by excluding pictures such as beds in top rooms or foods in bottom rooms (Palermo & Jenkins, 1964). With only two exceptions, children’s labels for the pictures conformed to adult labels (a 3-year-old child called a toy guitar a violin and a zebra a horse).

The following day, each child was shown 128 pictures, one at a time, randomly presented on plain white 6 in. $\times$ 3 in. (15.24 cm $\times$ 7.62 cm) unlined index cards. Half of these 128 pictures were the 64 pictures that were presented the previous day, and the other 64 pictures were novel toys, animals, and transports. For each picture, children were asked whether they remembered having seen it (“How about this one? Do you remember whether this was in one of the rooms we looked at yesterday? Was this something that you saw in one of the houses?”). After all 128 pictures were presented, children were given an unrelated distractor task that lasted 10 min, followed by a 5-min snack.

Then, for those pictures presented on Day 1 that had been recognized as well as those that had not been recognized, children were told,

> Now I am going to show you only pictures of things that really were in one of the rooms of the houses we saw the other day, OK? So, all of these next pictures are ones you saw in one of the rooms. I want you to point to the room in this house drawing where you think the picture was. This will be fun because we’ll see if you can remember which room each toy was in, OK?”

Children were shown a blank house drawing with its four rooms empty. If the child did not spontaneously point to a room in the house drawing upon being shown a picture, the experimenter probed each room in systematic order (rotating the starting room across children): “If you had to guess, do you think it was in this room? . . . This room?” The first room probed was systematically varied across children. Similarly, children were shown a white posterboard with empty houses in each of its four quadrants and asked to point to the quadrant in which the house containing that particular picture had been located (“If you had to guess, do you think it was in this house? . . . This house?”). The first house probed was also varied across children. Finally, children were shown the four colored posterboards and asked to point to the one that contained the picture, with the first color probed varied across children (“Do you think it was on this colored sheet? . . . This colored sheet?”). Thus, the order of probing for room, quadrant, and posterboard was systematized across children in a partial Latin square, whereas the order of probing each of the specific four options within room, quadrant, and posterboard was randomized.
Results

Across all analyses, an alpha level of .05 was used.

**Picture and attribute recognition.** Preliminary analyses established that the use of an alpha level of .05 was not significant difference whether the probing began with any particular room, house, or poster. Overall recognition levels to each combination of cues were unchanged by the order of questions. Table 1 shows the proportions of correct recognitions and false alarms in the picture recognition task by each age group and the proportion of correct recognitions for picture attributes. Correct recognition refers to the probability of accurate identification of a familiar picture or its attribute, whereas false alarm refers to incorrect identification of a novel picture as familiar. As can be seen in the table, younger children recognized fewer of the pictures than did older children (.64 vs .86) and correctly identified fewer of their attributes (.51 vs .71). Finally, younger children committed three times more false alarms to novel pictures (.58 vs .19). All of these values differed from chance values of either .5 or .25 (ps < .05, by two-tailed, Dunnett’s t test of multiple means against chance), and each of the age differences were statistically reliable.

On the basis of the false alarm (FA) and correct recognition (CR) data, we calculated two signal detection parameters to shed light on children’s recognition performance: (a) the probability (P) that a picture exceeds some threshold for accurate recognition (c), $P = CR - FA$, and (b) the level of certainty a child requires before deciding that a picture is familiar, signal detection parameter ($B_1 = FA/[1 - (CR - FA)]$). Lower values of $B_1$ indicate more stringent decision criteria. The ratio of correct recognitions/misses, false alarms, and the two signal detection parameters $P_i$ and $B_i$ were subjected to analyses of variance and regression analysis. The analyses of variance (ANOVA) included age (younger vs. older children) as a between-subjects variable and proximal versus distal contextual attribute level (room, quadrant, and posterboard) as a within-subject variable. As Table 2 shows, for nearly all dependent measures, there were significant main effects for age, $F(1, 104) = 9.53$, and cue level, $F(2, 104) = 8.41$, with mean square error between 12.23 and 26.50 ($p_{rep} = .96$). Only one of the significance tests for interactions was reliable ($P_{rep}$), the result of older children’s disproportionately greater discriminability ($P_i$) at the most distal, contextual attribute level (posterboard color). Regression analyses utilizing controls for differential success of proximal contextual attributes over distal ones resulted in findings that were consistent with these conclusions (see Appendix A).

**Relationship between picture recognition and localization.** In the next analysis, we examined the likelihood of recognizing the correct room, quadrant (house), and posterboard location of pictures that children previously claimed to have recognized, versus those they claimed not to have recognized. Data on how the recognition of pictures was related to the correct identification of contextual attributes appear in the top part of Table 2, both for pictures that were correctly recognized (+) and for those not recognized (−). The latter are instances in which children correctly placed a picture in its correct room, quadrant (house), and/or posterboard, even though they previously indicated not remembering having seen that picture. A given picture could be correctly

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>F Values Associated With Each Parameter in Experiments 1 and 2</strong></td>
</tr>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>CR</td>
</tr>
<tr>
<td>FA</td>
</tr>
<tr>
<td>$P_i$</td>
</tr>
<tr>
<td>$B_i$</td>
</tr>
<tr>
<td>CR</td>
</tr>
<tr>
<td>FA</td>
</tr>
<tr>
<td>$P_i$</td>
</tr>
<tr>
<td>$B_i$</td>
</tr>
</tbody>
</table>

Note. F values assessed with conservative degrees of freedom using the Huynh-Feldt correction formula (Winer, 1971). CR = correct recognitions; FA = false alarms; $P_i$ = signal detection parameter; $B_i$ = signal detection parameter.

$^1P_{rep} = .96$.

$^2P_{rep} = .88$. **$P_{rep} = .96$.**

ability to discriminate familiar pictures from novel pictures ($P_i$); their false alarm rate was significantly lower as was their $B_i$ parameter (suggesting a more stringent decision criterion). Planned contrasts revealed that younger children used significantly more lenient response thresholds on all three contextual attribute levels. Their discriminative sensitivity was also reliably lower than was older children’s on all three levels of the contextual attributes ($ps < .01$; $p_{rep} = .96$). Only one of the significance tests for interactions was reliable ($P_{rep}$), the result of older children’s disproportionately greater discriminability ($P_i$) at the most distal, contextual attribute level (posterboard color). Regression analyses utilizing controls for differential success of proximal contextual attributes over distal ones resulted in findings that were consistent with these conclusions (see Appendix A).

* $p < .05$. ** $p < .01$. *** $p < .001$. 

---

1. P; was used instead of the more common d', which can be derived from the z-score transformations of the above (CR and FA) values. This was done because the former indicators of discriminability and response bias, $Pr$ and $Br$, allow more fine-grained measurement when children respond at near chance levels, whereas traditional measures become less effective (Pollak, Cicchetti, Hornung, & Reed, 2000).

2. We used multiple ANOVAs rather than an omnibus multivariate analysis of variance (MANOVA) because the small sample size poses a more likely power problem for MANOVA than for ANOVA. Preliminary tests of sphericity revealed that homogeneity of variance assumption was easily met (ε values exceeding .90 in all cases).
Table 3
Proportion Accuracy for the Subset of Pictures Correctly Recognized (+) Versus not Previously Recognized (−) as a Function of Subsequent Attribute Accuracy in Experiment 1 and Prior Attribute Accuracy in Experiment 2

<table>
<thead>
<tr>
<th>Age group &amp; statistic</th>
<th>R</th>
<th>−</th>
<th>H</th>
<th>−</th>
<th>P</th>
<th>−</th>
<th>R + H</th>
<th>−</th>
<th>R + P</th>
<th>−</th>
<th>P + H</th>
<th>−</th>
<th>R + H + P</th>
<th>−</th>
</tr>
</thead>
<tbody>
<tr>
<td>Younger (52)</td>
<td>.70</td>
<td>.51</td>
<td>.62</td>
<td>.44</td>
<td>.37</td>
<td>.35</td>
<td>.77</td>
<td>.43</td>
<td>.53</td>
<td>.42</td>
<td>.46</td>
<td>.33</td>
<td>.27</td>
<td>.10</td>
</tr>
<tr>
<td>Older (38)</td>
<td>.84</td>
<td>.31</td>
<td>.85</td>
<td>.21</td>
<td>.61</td>
<td>.21</td>
<td>.90</td>
<td>.14</td>
<td>.81</td>
<td>.15</td>
<td>.69</td>
<td>.19</td>
<td>.55</td>
<td>.00</td>
</tr>
</tbody>
</table>

Note. R = room; H = house; P = poster.
*p < .05. **p < .01. ***p < .001.

There are several notable aspects of the data displayed in Table 3. First, older children generally possess more attributes in their memories for each recognized item, and this was true in this and the following experiment (see bottom half of Table 3). The differences are especially pronounced at the double- and triple-attribute level, with older children recognizing twice as many triple-attribute items (e.g., .55 vs. .27 in Experiment 1 and .65 vs. .29 in Experiment 2). This is important because 90%–100% of items that are associated with three attributes were correctly recognized; the problem is that younger children possessed significantly fewer of these. Second, younger children encoded more attributes for items that were unrecognized; for example, .33–.43 of items they failed to recognize were actually encoded with double attributes, a significantly higher proportion than was true for older children who only encoded double attributes for .14–.19 of unrecognized items. This suggests that they were less adept at retrieving items with the attributes in memory. One explanation for this retrieval inefficiency is that the memory traces of younger children are known to be less integrated (i.e., there is a weaker binding between the item and its attributes, so that recollection of the latter is less likely to evoke the former). Weaker binding would result in younger children’s worse performance in the positive (+) columns of Table 3 and better performance in the negative (−) columns, relative to older children.

Third, Table 3 does not reveal the number of items that were solely associated with a single attribute. For example, less than half of the .70 of R attributes were only recalled to R cues, the rest also being recallable to double- and triple-attributes; for older children, this was even more pronounced with over three quarters of the .84 items recognized in response to the R attribute also associated with the one or both of the other attributes. This is important to keep in mind lest one imagine that younger children had less attribute information in general available to recognize items. They had less double- and triple-attribute encodings but disproportionately more single attribute encodings than did older children. They had greater difficulty using their single and double (but not triple) encodings to recognize items. Consider that .33 to .43 of their double encodings were not recognized by them, which
was twice as high as the comparable proportions for older children’s failed recognitions (.14–.19).

Finally, one possible explanation of these data is that younger children were more random in answering the item recognition questions than the attribute source questions. However, there was no evidence that younger children applied stricter criteria in answering the attribute source questions than in answering the item questions; younger children were significantly more lenient in their response thresholds (B_{r}) on all three contextual attribute levels as well as for the item recognitions. In addition, as we show in Experiment 3, there is a very systematic relation between number of attributes retrieved and item recognition. In the absence of additional assumptions, these data are not consistent with an account in terms of random responding. If younger children were random in answering the item recognition questions, the same item accuracy should have been seen for all levels of attributes, but this was not the case as is seen in the following experiment.

Discussion

Replicating our pilot findings, 3-year-olds and 4-year-olds in this experiment frequently reported that previously encountered pictures were novel, even though they were later able to correctly identify attributes of these pictures’ contexts—the latter information being insufficient to lead to recognition awareness of the items themselves. This finding indicates that young preschool-aged children encoded information about events that for an older child were sufficient to generate a feeling of familiarity about having seen the item itself but that for younger children often was not.

It is important to note that younger children’s inability to recognize pictures that were associated with only one or two contextual attributes exceeded older children’s, usually by a factor of 2–3. In contrast, their inability to recognize the triple-cue pictures was relatively less pronounced (10%), though still significantly greater than that of older children (0%). This suggests a possible representational constraint on children’s recognition memory.

In the present context, a possible representational constraint is developmental differences in the types and amounts of information children encode about pictures. If the older children are more likely to integrate multiple aspects of a picture (semantic, location, color) then they may have a stronger basis for later recognizing previously presented stimuli—not because they are superior at monitoring their memories, per se, or because they are more likely to have encoded a picture’s individual attributes but rather because more potential retrieval cues exist in richly encoded traces to be accessed, and their traces are more richly encoded and integrated than are those of younger children, giving them greater retrieval support. This makes sense in view of the fact that younger children had far fewer pictures that were associated with all three attributes, but of those that they did have, their recognition was far better. Such an account implicating both encoding and retrieval differs from one that assumes younger children encoded similar information in memory but fail to introspectively examine it or, alternatively, that this information is somehow more opaque to introspection for them. Instead, the argument being suggested here is that the contextual attributes encoded by older children are more tightly integrated, so that retrieval of one will activate retrieval of the others, exceeding probabilities based on an independence of attribute assumption. Either possibility differs from an account in terms of inherent differences in metamnemonic awareness. Later, in Experiments 3 and 4 we examine these possibilities. But first, it is necessary to rule out an alternative explanation.

Experiment 2

In this experiment we address the possibility that the results of Experiment 1 are due to the relative weakness of the youngest children’s memory at the time of cuing its contextual attributes and that the act of cuing enhances later recall. This possibility has been raised in the classic encoding specificity studies of Tulving and his colleagues 3 decades ago. For example, Muter (1978), in an extension of Tulving’s discovery of words that could be recalled even though they had not been previously recognized, argued that “the recognition test may enhance performance on the later recall test” (p. 11). To address this possibility, we examined whether the strength of the encoded trace gets boosted by the recognition task itself. The reasoning is as follows: Perhaps younger children’s representations of the pictures at the time of the recognition task were too weak to support item recognition at that time. However, the mere presentation of the pictures as part of the recognition task may have been sufficient for the younger children to refresh their representations of the pictures, thus enabling them to subsequently use the contextual attributes to access the picture representations during the subsequent contextual cueing phase. Indeed, there is some evidence that preschool children’s verbatim memory traces are weaker than are older children’s, and the act of testing for recognition can lead to what has variously been called retrieval-time relearning or refreshing (Brainerd & Reyna, 1996).

If such an account is correct, then there is no need to posit representational constraints on metamemory as the basis of younger children’s seeming failure to recognize pictures whose attributes were identifiable in Experiment 1. This is because their poorer recognition performance can be due to weak initial memories for the pictures—until these pictures were shown again as part of the recognition task, at which point their traces were strengthened enough to be successfully cued later by the contextual attributes. There would be no need to invoke a deficit in metamemory that prevented them from recognizing the stored attributes and certainly no need to hypothesize that the richness and integration of the representation are driving this effect.

The procedure used in Experiment 2 was a modification of Experiment 1. The participants were presented the 64 pictures on Day 1. On Day 2, they were given the three context attribute probes and only later were they given the recognition test for the original 64 pictures. The sole change we made in the instructions for this experiment was the wording of the contextual cueing phase: Instead of telling children, as was done in Experiment 1, that the pictures presented had already been seen, we told them we were presenting pictures that may or may not have been seen. Included in this set of cued pictures were only the original 64 pictures that had been seen previously. If the presentation of the pictures in Phase 2 of the first experiment increased their trace strength, so that even when a picture was not recognized, its context subsequently was, then reversing the order of the context and item recognition tasks should result in fewer instances of failing to recognize items whose context had been. (In addition, the task reversal presents a conservative test of our hypothesis, given that the context recognition task in the first experiment entailed
only a single judgment—whether a given attribute was familiar. In contrast, in this experiment, because children were told that not all items had been presented previously, they had to essentially make two judgments in the context recognition task—item recognition and context recognition.)

Method

Participants. A total of 40 children were recruited for this experiment, 20 at each age, approximately half of whom were girls: 3-year-olds to 4-year-olds (M = 42 months, SD = 5.0 months) and 5-year-olds to 6-year-olds (M = 66 months, SD = 4.7 months). Four additional children, two at each age, were not included in the analyses because they were absent on Day 2. Children were recruited from the same regional day care centers, nursery schools, and kindergartens used for Experiment 1. Four of the younger children’s families were receiving Aid to Families with Dependent Children (welfare) payments at the time of the study, as was the family of one of the older children. None of these children had participated in the prior experiment.

Procedure. The procedures and stimuli were kept identical to those of Experiment 1, except for the reversal of the item and context recognition tasks on Day 2.

For the context recognition task, children were told,

Now I am going to show you some pictures. Only some of these were in the rooms of the houses, OK? If you think this is a picture you already saw in one of the rooms of the houses then I want you to point to the room in this house drawing where you think it was. But remember that not all of the pictures I am going to show were in the houses you saw, OK?

The original 64 pictures were randomly presented, and children were probed for each of their three contextual attributes, asking them to guess if they claimed not to know. Following this, these same original 64 pictures were interspersed among 64 novel pictures. The procedure from Experiment 1 was followed whenever a child did not spontaneously identify the location of the picture. If children said that they had not seen a picture, they were asked to guess its attributes.

Results

The bottom half of Table 1 shows the recognition accuracy for the 64 original pictures and the false alarms for the 64 novel pictures. As can be seen, although both age groups recognized items better than .5 chance (ps < .01, by two-tailed, Dunnett’s t test of multiple means against chance), older children once again recognized more pictures than did younger children, .93 vs. .78 (p < .01). Their performance ranged between 8% and 18% higher than in the first experiment, which was expected given that their recognition was tested after the contextual attributes had already cued the pictures rather than before, as was true in the first experiment. The data were subjected to a series of 2 (age: younger vs. older children) × 3 (contextual attribute: room, quadrant, and poster) ANOVAs. As in Experiment 1, the dependent variables for these analyses were correct recognitions, false alarms, and the two signal detection parameters P and B, described earlier. As can be seen in the bottom half of Table 2, for all four dependent variables, there continued to be significant main effects for age, Fs(1, 36) ≥ 10.00. Recognition accuracy and discriminability P, remained reliably greater at the more proximal level of cues (room and house) than at the most distal level of posterboard, with an effect of location cue, F(2, 72) ≥ 4.48, mean square errors between 6.01 and 16.84. As in Experiment 1, older children’s correct recognition rate was significantly higher than younger children’s (.92 vs. .78), as was their P. Younger children’s false alarm rate was again nearly 3 times higher than older children’s (.41 vs. .15, p < .01, prep = .96), as was their B parameter (p < .01, prep = .96). Planned contrasts confirmed that younger children used a significantly more lenient response threshold on the two most proximal cue locations, but no differences on this parameter emerged for the poster color attribute. Specifically, all but one of the age contrasts in Table 3 were reliable at the .05 level, with the exception being in Experiment 2, in which the percentage of pictures that were recognized even though their posterboard color had not been was equivalent for younger and older children (.33 vs. .30, p > .10).

As in the previous experiment, younger children’s discriminative sensitivity was significantly lower than was older children’s on all three contextual attribute levels (ps < .01; prep = .96). The sole divergence from Experiment 1 was that the interaction between age and contextual attribute level reached significance for correct recognition, with older children outperforming younger ones at the most distal level of contextual attributes (posterboard color). (P continued to remain reliable in the Age × Cue interaction.)

Because the designs of Experiments 1 and 2 were similar, the data from them can be combined into a single model to contrast the advantage of being given the picture recognition task after being given the context recognition task, as was done in this experiment. As expected, prior cueing with contextual attributes benefited later picture recognition—the mean proportion of pictures recognized by both age groups was higher in Experiment 2 than in Experiment 1, t(93) = 6.66, p < .001, rep = .99. Yet, even though overall recognition was higher for both groups of children in Experiment 2, failure to recognize pictures whose attributes were identifiable was still significantly more common for the younger group. As can be seen in the bottom half of Table 3, younger children correctly established one of the contextual attributes for 33% to 55% of the pictures they claimed not to recognize; they claimed not to have seen between 30% and 41% of pictures that had two of their contextual attributes established. Similar to the first experiment, 10% of pictures that failed to be recognized had nevertheless been previously associated with all three of their attributes. Once again, older children established the contextual attributes for far fewer pictures that they had failed to recognize (only 12%–16% of failed recognitions had been associated with correct contextual attributes), and they never failed to recognize pictures that had been previously associated with all three contextual cues.

Discussion

Identifying contextual attributes of pictures before testing for their recognition elevated recognition accuracy, as anticipated, but it did little to eliminate the failure to recognize pictures whose attributes were identified. Younger children continued to fail to recognize pictures they had previously been able to associate with their contexts. Therefore, the finding in Experiment 1 that young children fail to recognize pictures but subsequently correctly as-
sociate them with one or more of their three contextual attributes could not be due to refreshing the weak traces prior to cueing. Clearly, some contextual attributes were encoded by young children for these pictures and were available at the time of the recognition task in the first experiment, despite their inability to draw on this information to make familiarity judgments about the pictures themselves. Wais et al. (2008) found with adults that context recollection accuracy was significantly better than chance for items that were declared to be old and known, regardless of whether the context was presented before or after participants had just successfully recollected the item itself. The basis for age differences in the present experiment may involve both older children’s tendency to encode more tightly integrated, multiattribute traces and the subsequent advantage such traces convey during retrieval-time binding of attributes with each other and with the picture itself (Howe, 2000; Lloyd, Doiydum, & Newcombe, 2009). If true, this implicates both loci, though the present study was not designed to examine this issue. Consistent with past research, the more cohesive and integrated the information in the representation or trace, the better the performance is for the different elements contained in it (Anderson, 1991; Howe, 2000). In the final two experiments, we pursue this finding to elucidate the underlying mechanisms in monitoring and remembering differences between ages as well as within ages.

It is important to note that even the youngest children were able to recognize context information in the presence of the target item, that is, some (albeit impoverished) contextual cues. They also did better recognizing targets when cues were presented prior to the item recognition test rather than after it. This suggests that items and context are not well integrated in younger children’s memory, an idea consistent with several other theories (trace integration theory, information integration theory, and memory binding) that indicate that attributes are less tightly connected to the item itself as well as to each other (Starns et al., 2008).

Experiment 3

Experiments 1 and 2 (as well as the pilot experiment) revealed pronounced age differences in the failure to recognize items whose attributes had been previously recognized: 3-year-olds and 4-year-olds were significantly more likely to fail to recognize pictures about which they could access (and in Experiment 2, already had accessed) context information. This failure cannot be due to any sort of refreshing or relearning process that occurs around the time of the recognition task because even when the recognition phase follows successful cueing, there remain large age differences in failing to recognize pictures that were correctly associated with one or more of their three contextual attributes.

In Experiment 3, we examine the relationship between picture recognition and representation of its context. We ask, specifically, whether the observed developmental trend in the failure to recognize items whose attributes had been previously recognized is the result of unequal trace representations for the pictures of younger and older preschoolers. If it turns out that developmental differences in recognition memory fade when both age groups encode similar amounts of contextual information then this would implicate representational factors as the basis of developmental differences in the failure to recognize items whose attributes had been previously recognized. Perhaps far fewer of younger children’s traces contain multiple contextual attributes of the original picture. Thus, even though younger children retrieve fewer pictures than do older children, younger children’s recognition accuracy is nevertheless much higher when pictures of items are associated with all three attributes of room location, house location, and posterboard color (90%).

Earlier experiments with adults have shown that the encoding of contextual attributes has a nonlinear, multiplicative effect on retrieval. Multiple attribute recognition of a given item is associated with higher likelihood of recognition of the item than would be predicted based on the sum of the probabilities of the individual attributes successfully prompting its recognition (Jones, 1976). A similar nonlinearity has also been shown to exist for 7-year-old to 10-year-old children (Ceci, Lea, & Howe, 1990). The failure to recognize items whose attributes had been previously recognized by the younger children in Experiments 1 and 2 thus could be due to their encoding fewer contextual picture attributes (and fewer combinations of them) than the older children encoded. In the present experiment, we attempted to examine with greater specificity whether there are age differences in the likelihood of recognizing items for which all three contextual attributes were retrieved versus items for which fewer contextual attributes were retrieved.

To address this question, we used a procedure similar to Experiment 2. The only difference was that because of time constraints, the contextual attribute cuing and the recognition tasks were presented on separate days. However, the analysis focused on the independence of the attributes and, particularly, on whether an accurate response to one attribute would raise the probability of an accurate response to another. Further, we were interested in whether this would occur not only between-subjects but also within-subjects.

Method

Participants. Forty-eight children from backgrounds similar to those in the prior experiments participated, 24 at each age, with roughly equal numbers of each sex. However, 3 of the younger and 2 of the older children were unable to finish the tasks as a result of a class activity, and 6 others were absent from one or more sessions. In the end, 18 younger children ($M = 41$ months, $SD = 5.2$ months) and 19 older children ($M = 65.7$ months, $SD = 6.6$ months) participated in all 3 days of the experiment. One of the younger children and 1 of the older ones were receiving free or reduced lunch.

Procedure. Children were tested individually with the same set of stimuli as used in the previous studies and a procedure similar to that of Experiment 2. The difference was that the recognition task was administered on a separate day, Day 3.

Results

In the study, we aimed to assess the role and extent of age differences in the likelihood that successful recognition of one attribute (i.e., room location, house location, posterboard color) leads to subsequent successful recognition of the other two attributes. We also aimed to assess whether the recognition of items correctly associated with all three contextual attributes was greater than would be expected on the basis of the recognition of items correctly associated with only one or two attributes. We focus the presentation of the results around these main goals.
As can be seen in Table 4, older children’s recognition was disproportionately made up of triple-attribute-cued pictures (54%), whereas only 34% of younger children’s recognitions comprised triple-attribute-cued items (p < .01). Twenty-seven to 31% of younger children’s Day 3 recognitions had previously been accessed by one or two attribute cues on Day 2 (vs. 15%–29% of older children). Younger children were able to recognize 8% of the items that were unsuccessfully cued by any single attribute (vs. 2% for older children, a difference that failed to reach significance p > .10). Thus, many more of the older children’s item encodings contained multiple attributes than was true of younger children, and older children’s recognitions were disproportionately associated with multiple correct cuing.

Table 5 presents the flip side of the above data, namely, the proportion of pictures that were successfully associated with zero, one, two, or three attributes on Day 2, which were subsequently recognized on Day 3. If the recognition process is random, any given picture has a .5 chance of being correctly recognized. However, as can be seen from the data in Table 5, only 14% to 18% of the original pictures for which children recognized zero attributes were subsequently recognized on Day 3. Clearly, neither age group simply guessed at random about the triple- and double-encoded items.

Table 5 also shows that younger children lag behind their older peers in the use of single and double attribute traces to support recognition. Only 37% of single contextual traces were later recognized by young children, compared with 66% of such items for older children. Likewise, only 68% of pictures that were accessed by two contextual cues were later recognized by younger children, compared with 91% for older children (ps < .01; p<sub>rep</sub> = .95). It was only in the case of items that had been accessed by all three contextual attributes that younger children’s recognition accuracy approached the perfect levels of older children. Younger children correctly recognized 97% of the pictures that were retrievable to all three cues on Day 2, and older children correctly recognized 100% of these triple-cued pictures, t(36) = .25, p > .10; p<sub>rep</sub> = .61.

Next, we tested the hypothesis that age differences in the likelihood with which prior successful cuing with one contextual attribute leads to subsequent successful cuing with the other two contextual attributes. We used the algorithms developed by Ceci et al. (1980) to calculate the valences for each attribute’s probability of cueing correct recognition for a picture, with R, Q, and P representing the single valences for the room, quadrant, and poster, respectively. The conjoint probability of cueing correct recognition of a picture by cueing with its room and house quadrant V(RQP) corresponds to the observed recognition valences for the single

<table>
<thead>
<tr>
<th>Age group &amp; statistic</th>
<th>Zero</th>
<th>Single</th>
<th>Double</th>
<th>Triple</th>
</tr>
</thead>
<tbody>
<tr>
<td>Younger</td>
<td>.08</td>
<td>.27</td>
<td>.31</td>
<td>.34</td>
</tr>
<tr>
<td>Older</td>
<td>.02</td>
<td>.15</td>
<td>.29</td>
<td>.54</td>
</tr>
<tr>
<td>t(47)</td>
<td>0.70</td>
<td>2.19*</td>
<td>0.11</td>
<td>3.16**</td>
</tr>
</tbody>
</table>

* p < .05. ** p < .01.

When the valence estimates for each child were subjected to one-way ANOVAs for the single, double, and triple valences, the result showed significantly more triple valences for the older children only, F(1, 41) = 7.32, MSE = .011 (p<sub>rep</sub> = .95), ƞ<sup>2</sup> = .221 (power = .781) but no age differences for single or double valences (p<sub>rep</sub>'s < .65). This indicates that for older children, there were significantly more pictures simultaneously cued with all three contextual attributes. It is critical to note that whereas 100% of the older children’s valences for the triple cue exceeded the estimates derivable from the successes of their three independent single cues, this was the case for only 42% of the younger children’s triple valences (p < .05; p<sub>rep</sub> = .88), by two-tailed t tests. Although it is not possible to compare the 42% of younger children’s triple-cued valences with 100% of older children’s in terms of the independence hypothesis because, by definition, all 42% exceeded it as did all 100% of the older children’s, the advantage of triple cued items was evident within subjects: When we restricted the analysis to the subset of 42% of the younger children’s triple cued items, these cases behaved similarly to the 100% of older children’s triple cued items.

Hence, younger children’s inferior recognition overall was associated with (a) their relative failure to recognize single- and double-cued items, (b) their encoding of relatively fewer triple-cued items, even though they recognized these items with accuracy equivalent to that of older children when the younger children did encode the items, and (c) greater statistical dependence among the three levels of contextual cuing for older children, indicating that if one cue was successful, the odds were elevated that the others would also be successful. In contrast, the younger children often behave as if the success of one contextual cue was independent of the others. In other words, if older children knew the room location of an item, they were more likely to also know its house location and posterboard color. This was not routinely the case for younger children, for whom knowledge of one attribute did not automatically mean greater likelihood of knowledge of the other two attributes.

**Discussion**

In this experiment, we examined the hypothesis that the recognition process and the failure of younger children to recognize items whose attributes had been previously recognized may benefit...
from the joint activation of multiple contextual attributes in an item’s representation, especially if retrieval of single and double-
attribute traces is ineffective as was true of younger children. Even
though all memory representations had been derived from exactly
the same stimulus presentation, age differences were apparent in
the amount of the contextual attributes that children spontaneously
encoded, with older children encoding more triple-contextual-
attribute information. If older children recognized a picture at all,
they were more likely to identify all three of its contextual at-
tributes than to identify only one or two of them. This noninde-
pendence of picture recognition and multiplicity of cue successes
was significantly less likely to be true for younger children. Esti-
mates of the triple-cue successes for younger children did not
exceed the estimates based on the summation of the independent
cue valences. They performed as if their response to one cue had
no consequence for responses to the other cues. In other words, for
young children, knowing any one attribute was entirely indepen-
dent of knowing the other two attributes.

It is noteworthy that encoding all three contextual attributes of
a picture led to its nearly invariant correct recognition at both ages.
It was only in this situation that failing to recognize pictures whose
attributes were recognized disappeared for younger children. Devel-
opmental differences in recognition failure were particularly
pronounced when pictures had been retrievable with only one or
two attribute cues. Older children were significantly more likely
than were younger children to recognize a picture that had been
accessible by one or two contextual cues. In contrast, younger
children failed to recognize between one third and two thirds of all
pictures for which they had retrieved only one and two contextual
attributes. Thus, younger children’s representation of the encoded
contextual attributes was less tightly integrated, making the cue
information less effective in arousing the contents of the picture’s
entire record. Consequently, the observed deficit in preschoolers’
recognition performance in Experiments 1 and 2 may have been the
result of a reduced amount of information in their memory
representations (i.e., fewer triple-cue valences) rather than an
inability to monitor the contents of their memory. If the latter was
the cause, then their success on the triple-cued items would not
have been nearly perfect because they would still have faced the
task of judging whether the items were familiar which, if
metamemory was undeveloped, would have introduced the same
amount of error observed for items with fewer attributes. Clearly,
they were very effective at monitoring triple-cue items, and their
problem was due to their relative lack of such items rather than a
monitoring deficit. Alternatively, one could argue that younger
children may have encoded as many contextual attributes as older
children (and as many triple-cue items) but may have forgotten
these attributes more rapidly. If true, the current pattern of results
could be reinterpreted in terms forgetting rather than in terms of
canonical age trends in encoding rich representations that are more
readily retrievable. Although this is possible, prior research using
similar paradigms has revealed that 7-year-olds are superior to
4-year-olds even after only 20 s delay (Ceci & Howe, 1978a,
1978b; Ceci et al., 1980), rendering a forgetting explanation un-
likely over such brief intervals.

Are these results due to an interaction between metamemory and
a picture’s mental representation, such that relatively rich repre-
sentations of pictures—ones that contain all three of the relevant
contextual attributes—enable metamemory to work more effi-
ciently than representations containing fewer attributes? If true,
this leads to the predictions that young children possess the ability
to track and monitor their memory and that this ability is mediated
by the richness of the representation. In situations in which young
children’s representations are as rich as older children’s in terms of
the amount of contextual attributes that were encoded, their mon-
toring should be equivalent and their failure to recognize items
whose attributes had been previously recognized should be simi-
larly rare, a prediction we test in the final experiment. Considering
the three main findings together suggests that both encoding- and
retrieval-time processes may be involved in the age trends in
recognition failure: When the items were triple encoded there
were no age differences in recognition, thus supporting the
importance of encoding rich representations; however, younger
children were far less effective at retrieving items when only
single and double cues were encoded, so to some extent encod-
ing constrains retrieval possibilities.

In sum, as we predicted, recognition memory was associated
with the quality of children’s representations (Ceci et al., 1980;
Jones, 1976). Thus, it seems possible that metamemory operates
less efficiently when representations are impoverished. That is,
representational quality may mediate the effect of metamemory on
recognition memory. In the final experiment, we use a novel
analytic approach and a test for mediation to directly test the
relation among metamemory, recognition memory, and represen-
tational constraints.

**Experiment 4**

Experiment 3 suggests that recognition failure of items whose
attributes can be recognized occurs primarily when the memory
representation is impoverished (i.e., comprising fewer than triple-
attribute encodings). When younger children’s memories were
represented by triple-attribute encodings, such failure all but dis-
appeared. Hence, there does not seem to be a fundamental
metamemory deficit per se in young preschoolers, but rather there
does seem to be an abundance of relatively impoverished repre-
sentations, that is, those containing limited contextual information,
which perhaps constrains monitoring. Thus, developmental
changes in recognition memory and metamemory could be af-
fected by the nature of the representations available (Ornstein &
Elishberger, 2004).

In this final experiment we examined this hypothesis directly.
To provide generality, we used an independent means of opera-
tionalizing a memory representation, to triangulate the valence
evidence from Experiment 3. To test the role of representational
constraints on monitoring directly, we also included a modified
metamemory task, and we used different (longer) retention inter-
vals, stories rather than pictures, and older children to provide a
daring prediction—that preschoolers would monitor as effectively
as 9-year-olds when their representations were similar, an even
larger developmental span than was used in the prior experiments.
We endeavored to test whether metamemory mediates the effect of
representational constraints on recognition memory. This experi-
ment involved several phases to operationalize children’s repre-
sentations, one of which is reported in Ceci et al. (2007), and we
use their representational values as predictors in this experiment.
The metamemory data and the mediation analyses involving it are
unique to this experiment.
The first part of the study was a stimulus judgment task used by Ceci et al. (2007) to derive representational estimates, and the second part was a metamemory and recognition memory task. In Ceci et al.’s (2007) stimulus judgment task, children sorted triads of common objects (animals, foods, etc.), indicating which object in the triad did not belong with the other two. This task is similar to the popular Sesame Street activity, “Which of these three is not like the others?” By scaling children’s answers to these triads, Ceci et al. (2007) operationalized aspects of children’s semantic representations and, it is important to note, manipulated them to make a priori predictions about the relationship between this aspect of children’s representation and their suggestibility. We adopt this approach here and make a priori predictions about the relationship between this aspect of children’s representation and their metamemory performance, using the representational values reported by Ceci et al. (2007). Thus, Experiment 4 was designed to pose what for developmental psychologists is the single most powerful form of prediction, namely, that of conditions that lead to reverse developmental outcomes (Brainard et al., 2008). In an extension of Ceci et al. (2007)’s results, we examined the role of representational space in mediating the relation between metamemory efficiency and recognition memory. Because the scaling data were only available for 4-year-olds and 9-year-olds, in Experiment 4 we used this age range rather than 3-year-olds and 6-year-olds. However, for our purposes, this age range is preferable because 4-year-olds are known to possess immature metamemory, whereas a full-fledged understanding of memory appears to be fairly developed by the end of middle childhood, with 9-year-olds possessing substantial metamemory (Jaswal & Dodson, 2009). Also, as noted, this age range extension was intended to increase generalization, along with longer retention intervals, different operationalization of representations, and story stimuli rather than pictures.

In Experiments 1–3, memory representations were operationalized in terms of the number of contextual attributes and the likelihood that retrieving one attribute elevated the odds of retrieving other attributes (Experiment 3). In this experiment, memory representation is operationalized in terms of the semantic distance between items, in an attempt to broaden the evidentiary base for earlier conclusions about the critical role that representations play. The Euclidean distance between two items in semantic space is an aspect of their representation, with zero space corresponding to a lack of differentiation between them (total attribute overlap), compared with when the same items are represented farther apart as part of a dense semantic network of independent but integrated attributes (Ceci et al., 2007).

Remember–know judgments are measured by answering questions, some of which are inappropriate for preschoolers, such as “Do you remember what you were thinking when the item was presented?” (Tulving, 1985). The metamemory task in this study consisted of a child-friendly modification of the typical remember–know judgment task, and all of the children appeared to understand and respond appropriately to the changes we made. Remember judgments are assumed to reflect greater certainty based on encoding and monitoring processes than are know judgments, which tend to be vague intuitions lacking episodic details (Nelson & Narens, 1990).

Of interest is whether remember–know judgments are influenced by the nature of the mental representation for the items depicted in a story presented to children. Remember judgments are accomplished by a recollective process involving the retrieval of episodic details from an earlier encoded experience (Nelson, 1996; Yonelinas, 2002). Stahl and Klauer (2009) have shown that remember judgments are saturated by verbatim trace information, whereas know judgments are more gistlike.

The data from the three preceding experiments suggest that if younger children’s impoverished representations are constraining their ability to monitor pictures that they encoded, then this deficit should disappear in situations in which their representations were similar to those of older children, as suggested by the triple-cue data in Experiment 3. The reverse developmental prediction is that when older children’s representation is more impoverished than is younger children’s, they should exhibit less accurate metamemory performance. Therefore, if inefficient metamemory is behind the failure to recognize items whose attributes have been correctly identified, when older children’s representation is more impoverished than younger children’s, older children should make more recognition errors of items that were previously cued by attributes than should younger children. Again, these predictions derive from the hypothesis that although metamemory failure is involved in this phenomenon, it is mediated by the nature of children’s representation. If younger children use metamemory effectively when their representations are equivalent to those of older children, then this would be evidence that developmental differences in recognition accuracy in general, and in failing to recognize items whose attributes have previously been recognized specifically, cannot be attributed to a global inability to use metamemory or to inadequate neural systems to support monitoring.

Specifically, it was expected that in their metamemory judgments, children would accurately distinguish between critical items and distractors when they were distally represented, but not nearly as accurately when items were closely represented. This is because the semantic attributes of closely represented items overlap, and therefore, encoding of the critical item activates some of the nearby items’ representations, making memory monitoring more difficult. In other words, a child exhibiting good metamemory should be more likely to respond with either a reject or a know judgment to a semantically proximal distractor because even though there are no episodic markers for it in memory, its gist would have been partially encoded as a consequence of encoding a related critical item. In contrast, encoding a critical item in the story activates very little of the representation of a distal distractor, thus making it easier to distinguish between these items during monitoring. By this reasoning, the key prediction would be that although the proximally represented distractors would be difficult to monitor, age differences in metamemory would be reduced for items with distal distractors, that is, younger and older children would be more similar at making accurate remember judgments of the original critical items and rejection of their distractors when the latter are distally represented. The approach taken here allows judgments of what is proximal and distal to be individualized for children, so that what is proximal for one child may not be for another, leading to targeted predictions, including reverse developmental trends when the same pair of items are distally represented for younger children but proximally represented for older children and vice versa.
Method

Participants. Seventy-four children participated in this study. Half were 4-year-olds and half were 9-year-olds ($M = 57$ months, $SD = 2.2$ months; $M = 116$ months, $SD = 6.2$ months). About half of the participants were female. The children came from schools and preschools serving working class and middle class families.

Materials. Materials published by Ceci et al. (2007) were used as stimuli, as they had been scaled so that the semantic distance between them was known. Here we describe the procedures used by those investigators before describing how we utilized their materials to test our metamemory hypotheses.

Six younger children and six older children in Ceci et al. (2007) rated the similarity of triads of pictures from a set of 22 commercial drawings of familiar foods, insects, and animals. The 22 pictures were of milk, soda, cheese, cake, lemon, watermelon, egg, deer, cow, robin, lobster, orange, ant, squirrel, eagle, bear, pig, grapefruit, spider, butter, horse, and crab. First, the children were directed to name each picture as the experimenter pointed to it. Then, each child was presented with 257 triads, systematically assembled from rating every 6th triad in the complete set of all permutations of three of the 22 pictures:

$$\frac{22!}{(22 - 3)!3!} = 1,540, \quad \frac{1,540}{6} = 256.6 \quad (2)$$

On each trial, children were presented with three pictures and asked, “Which of these pictures does not belong with the other two?” After indicating which picture was unlike the others, children were asked to explain their choice. (The 257 triads were administered by Ceci et al., 2007, to each child individually over a period of 3 to 5 weeks; this permitted the computation of age level Euclidean distances between items without requiring each child to judge every one of the permutations.)

Similarity ratings of the 22 items were entered into individual $22 \times 22$ matrices that were subjected to nonmetric multidimensional scaling (ALSCAL). Separate INDSCAL models were run for each child to produce individual solutions that represented the relative distances between each of the 22 items in semantic space for that child. The proximity between each critical item and its distractor for each child was assessed by their Euclidean distances and operationalized here as an aspect of their representations.

On the basis of the multidimensional scaling, 8 pairs of objects were chosen for the recognition memory task. An object from each pair was assigned to one of two lists, A or B. Children were presented with an illustrated story about a trip to the zoo (see below) that contained pictures of the eight objects from one of the lists (critical items), as well as pictures of another eight objects that were not mentioned in the story and were not semantically related to the critical items. In addition, eight items that were semantically related to the critical items from the list but were not used in the story served as distractors in a subsequent recognition task. Thus, two versions of the zoo story were created by swapping critical items and distractors so that the distractors in one list served as the critical items in the other list.

The story described two children’s visit to a zoo. Depending on which version they were assigned to, children saw an eagle (or a robin) in the aviary, they drank a glass of cold milk (or soda), they ate a cheese (or egg) sandwich at the snack bar, they rested in the shade of a lemon (or orange) tree, they saw a spider (or ant) in the house of insects, they watched a horse (or bear) in a breeding area, they saw a lobster (or crab) in the aquarium, and they watched a cow (or deer) in a field on the bus ride home. Hence, one member of each pair was the critical item that actually appeared in the story, and its semantically related pair member was the distractor for the recognition test. As the story was presented, the eight critical items and the eight unrelated noncritical pictures were set on a large posterboard, each placed in a room in one of four 4-room houses as in Experiments 1–3. By the end of the story, all 16 of these items occupied a room in one the four houses. Another 16 pictures of objects, not rated in the stimulus judgment task, were used in the recognition task as foils, as well as the eight semantically related distractor items that had not appeared in the story.

Procedure. Children were tested individually. Two days after they were presented the zoo story, they were interviewed about it. At this point, half of the children were given misinformation about the original critical items in the story. Specifically, it was suggested that the distractor items had been part of the story rather than the actual critical items. For example, children were told by the interviewer, “Wasn’t that an interesting story about Jessie’s trip to the zoo when she and her brother saw that eagle while they drank their sodas?” when in fact in that version of the story, Jessie and her brother had seen a robin, not an eagle, and they had drunk milk, not soda. As noted, the critical items and their suggested distractors were flipped, so that the critical items for one child’s recognition test were the distractors for the other child, and vice versa, to control for item-specific effects. The remaining half of the children were interviewed but given no semantic misinformation. This was done to assess how difficult it would be to distinguish the critical items from their distractors in the absence of erroneous suggestions about them. As there were no differences between these conditions (they did not interact with age or distance or Age $\times$ Distance in the prediction of metamemory scores; $ps > .10$), we shall collapse across them in discussing the results for this experiment.

Five to 7 days after the interview (7 to 9 days following the story presentation), children were presented with 32 pictures in a standard serial recognition test and asked to decide for each picture whether it had been part of the story. The test comprised the eight critical items from the story, eight distractors that were not part of the zoo story, and 16 novel foils. The order of presentation of these 32 items was randomized for each subject.

The metamemory task immediately followed the recognition task and comprised only 16 items, the eight original items from the zoo story and the eight distractors. These items were presented randomly to each child. Children were asked to make either a blue or red crayon mark on each item, corresponding to a know–remember judgment about it, with half told to use blue to indicate they remember and the others told to use red for this judgment.

Children were instructed to make a mark (red or blue, depending on the counterbalancing) if they had episodic details still in memory (e.g., could recall what house it was in, what object it was next to) and if they had definite memory of its appearance (i.e., if you “really, really remember seeing the picture in the zoo story”). Children were instructed to make a differently colored mark for items that they felt they knew from the story even though could not actually recall seeing it and could not recall anything about its occurrence or appearance in the story or what it was next to but
nevertheless felt that it had appeared in the story. ("In other words, if you don’t really, really remember having seen it clearly but you just know it was in the story then make a blue mark on it").

Each answer was classified as remember or know based on the child’s red or blue mark but also the child’s verbal responses to the interviewer’s probes about its episodic features. The interviewer categorized each child’s answer in terms of remember–know–indeterminate, and a second rater coded 10 of the children’s answers to perfect agreement. The feeling of knowing (as opposed to remembering) that a distractor item was in the story is an indicator of metamemorial awareness because in the absence of explicit episodic memory of its appearance, good metamemory would be reflected in either know or reject rather than claiming to remember. This is because, even though the distractor item’s gist is related to the critical item’s gist, there is no episodic part of its representation to warrant an explicit remember judgment. If children mistakenly claim to remember distractor items, this would be evidence of poor metamemory, as would claiming to merely know critical items as opposed to explicitly remembering them. Of interest is whether such metamememonic awareness varies with an item’s representation, here operationalized as the Euclidean distances between critical items and their distractors.

Results

Relationship between metamemory and representational space. Each pair of critical-distractor items was analyzed as a function of how closely its two members were represented to each other. As noted, a child’s correct metamemory judgment was based on indicating that they remembered the critical item in a pair (because giving a critical item a know judgment would indicate a lack of certainty and failure of a recollective process that would be warranted by the item’s episodic details), coupled with either rejecting outright or indicating a know judgment for its distractor. (A remember response to both the critical item and its distractor item was not taken as evidence of metamemorial insight, nor was a reject or know judgment to both items.) Although Euclidean distances between critical items and their distractors were different for the younger and older children, we expected that independently of age, false recognition of distractors would be higher when they were represented closer to critical items than when they were represented further from them.

We used the coordinates from Ceci et al.’s (2007) INDSCAL (Carrol & Chang, 1970) solutions to calculate Euclidean distances between each pair of critical-distractor items for each age. Out of seven pairs for which younger and older children differed in their representational distances, four were ones that older children represented closer together than did younger children, and the remaining three pairs were ones that younger children represented closer together than did older children. Table 6 lists these pairs. The far right column presents the predicted developmental differences in metamemory scores (i.e., accurate remember–know judgments).

The pairs for which we predicted that younger children would show higher metamemory memory accuracy and the pairs for which we predicted older children would show greater accuracy were grouped. Next, we calculated the proportion of pairs for which each child exhibited accurate know–remember judgments. A child had to answer both pair members correctly to be part of the proportion correct—a conservative criterion. This meant that the three pairs for which we predicted that younger children would be superior in their monitoring would be based on the proportion of the six items (three pairs) they answered correctly, and the four pairs for older children’s superiority would be based on the proportion of eight items answered correctly.

The mean untransformed proportions of correct remember judgments as a function of age and representational proximity (repre-

### Table 6

<table>
<thead>
<tr>
<th>Pair</th>
<th>Distance</th>
<th>Younger</th>
<th>Older</th>
<th>Predicted metamemory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lemon–Orange</td>
<td>1.31</td>
<td>0.89</td>
<td>Younger &lt; Older</td>
<td></td>
</tr>
<tr>
<td>Egg–Cheese</td>
<td>1.61</td>
<td>0.23</td>
<td>Younger &lt; Older</td>
<td></td>
</tr>
<tr>
<td>Cow–Deer</td>
<td>0.92</td>
<td>0.57</td>
<td>Younger &lt; Older</td>
<td></td>
</tr>
<tr>
<td>Ant–Spider</td>
<td>0.36</td>
<td>0.32</td>
<td>Younger = Older</td>
<td></td>
</tr>
<tr>
<td>Milk–Soda</td>
<td>0.72</td>
<td>1.66</td>
<td>Younger &gt; Older</td>
<td></td>
</tr>
<tr>
<td>Robin–Eagle</td>
<td>0.50</td>
<td>1.47</td>
<td>Younger &gt; Older</td>
<td></td>
</tr>
<tr>
<td>Bear–Horse</td>
<td>1.07</td>
<td>1.93</td>
<td>Younger &gt; Older</td>
<td></td>
</tr>
<tr>
<td>Crab–Lobster</td>
<td>0.38</td>
<td>0.70</td>
<td>Younger &gt; Older</td>
<td></td>
</tr>
</tbody>
</table>

Note. < is less confusable; > is more confusable; = is equally confusable. INDSCAL =.

---

3 Remember–know judgments were developed for use with adult participants and depend on distinctions between awareness of episodic details being retrieved (e.g., thoughts, imagery, or associations induced by an item at the time of study that a participant can later recollect) and vague feelings of semantic memory output in the absence of explicit images or associations that prompt a know judgment because they lack episodic details. Instructions for making traditional remember–know directions are beyond the understanding of young children. We adapted the procedure Billingsley Smith, and McAndrews (2002) used successfully with 8-year-olds who were instructed that remember judgments needed to have actual details recalled in contrast to know judgments in which they were confident that they had encountered a picture but could not recall any particular detail, such as its location. In training, we provided children with several of the actual statements made by adult participants to describe the basis of their–remember–know differences; children were given six of the reasons given by adult participants for remember versus know decisions in Gardiner, Ramponi, and Richardson-Klavehn (1998). For example, children were informed that if they remembered seeing an item because it had appeared next to another item they remember seeing, then they should use the color that corresponded to remember (e.g., remembering they saw x because they remember that it was next to y). On the other hand, children were instructed that if they did not remember actually seeing an item but know it was there, then they should use the other color (e.g., “Use this color if you are pretty sure you saw x but you can’t actually remember seeing it”). During training, we asked children which color they would use under conditions of remember and know, such as, “Suppose a child remembers seeing a sofa in one of the rooms because she remembered thinking about a child in her class named Sophia when she first saw that picture of the sofa. What color should she use?” Versus “Suppose a child says to herself, ‘I am sure I saw this picture but I can’t remember why I think I saw it’. What color should she use?” Six of these distinctions were adapted from the actual transcripts in Gardiner et al. (1998) so that they would be understandable for young children. We worked with children on each contrast until they appeared to understand the distinction before we gave them the actual remember–know questions used to derive their metamemory score.
sented closer vs. further apart) are displayed in Figure 3. There was no effect of item list, so the data are combined across both versions. As can be seen, both age groups were roughly twice as accurate indicating remember judgments for critical items that were distally represented than they were for items that were proximally represented. Younger children correctly indicated 69% of items as remember when they were represented more distally by older children, and older children correctly indicated 61% as remember when they were more distally represented by younger children than by their own age group. Both age groups exhibited very low false alarms for items that it represented distally (overall 9% versus 23% for items that it represented proximally), and there were no age differences in either of these averages ($p > .10$). Thus, the manner in which information is represented influenced age trends not only for memory but also for monitoring.

A $2 \times 2$ (age group: proximally represented vs. distally represented) repeated-measures ANOVA was conducted on the arcsine-root transformed proportions of know–remember judgments. The omnibus analysis revealed an Item-Group × Age interaction, $F(1, 70) = 31.10, p < .001$, $\eta_p^2 = .31$. When the items were represented closer for older than for younger children, the younger ones correctly judged a greater proportion of the original critical items as remembered and either rejected or rated as know these items’ distractors, $F(1, 70) = 71.73, p < .001$, $\eta_p^2 = .51$. When the items were distally represented than they were for items that were represented proximally, older children correctly indicated 69% versus 23% for items that it represented proximally), and there were no age differences in either of these averages ($p > .10$).

Mediation. A meditational analysis of the role of representations. This would directly test whether the failure to recognize items whose attributes were recognized by younger children in Experiments 1–2 and the relation between metamemory and recognition memory in Experiment 4 were both mediated by representational quality (Euclidean distance and number of attributes). It is important to note that it would allow this determination after taking into account the influence of PFC maturity (age). We conducted several analyses to test the hypothesis that children’s mental representation, as reflected in Euclidean distance, mediates the influence of metamemory on recognition accuracy. These models were run with and without age entered as a covariate. Following Frazier, Tix, and Baron (2004), we first established associations between each of the variables (the independent variable being metamemory and Euclidean distance as the mediator) with the accuracy of recognition memory as the dependent variable. Next, we regressed recognition accuracy on both the representational and metamemory variables, statistically removing their influence by using the residuals from the regression analyses. The independent variable (metamemory) and mediator (Euclidean distance) each played a significant role on recognition memory accuracy when considered independently, with Sobel $z$ values ($p < .05$; Preacher & Leonardelli, 2001).

We conducted three more regression analyses. First, the dependent variable (recognition accuracy) was regressed on the independent variable (metamemory), resulting in the coefficient corresponding to a path between them. In the second analysis, the mediator variable was regressed on metamemory to obtain the regression coefficient for the path between them ($B = 0.160$). In the third analysis, the dependent variable (recognition accuracy) was regressed simultaneously on both the mediator and the independent variable. This produced unstandardized regression coefficients for the paths between the independent variables and mediator variables.

For mediation by Euclidean distance to be able to explain the observed findings, the following would logically need to be true:

- Metamemory accuracy would predict the dependent variable, recognition accuracy.
- The mediator variable, Euclidean distance, would predict metamemory accuracy (or vice versa).
- Euclidean distance would predict recognition accuracy.

![Figure 3](image-url) Mean metamemory judgments by younger and older children for items that are represented by peers as closer together or further apart for each age in Experiment 4.
• When Euclidean distance is added to the initial model, the standardized coefficients for metamemory predictiveness of recognition accuracy should decrease significantly.

The first series of regressions assessed the magnitude of direct and indirect effects when age was versus was not controlled. Metamemory accuracy was first regressed on children’s recognition accuracy, yielding significant results, ($\beta_s = -.545$ and $.593$; $p < .001$). Metamemory accuracy predicted children’s Euclidean distance ($\beta_s = .771$ and $.611$; $p < .001$). Finally, with both metamemory accuracy and Euclidean distance in the model, the standardized coefficient of the former was reduced to .102 ($p > .10$), whereas Euclidean distance remained a strong predictor of recognition accuracy ($p < .01$). The total effect of Euclidean distance was computed by summing the direct and indirect effects, using the standardized coefficients. The indirect effect was .17 (calculated by multiplying the two corresponding paths linking Euclidean distance, $.67 \times .25$). Unstandardized regression coefficients and standard errors for paths between metamemory and Euclidean distance and between Euclidean distance and recognition accuracy were the basis of the $Z$ statistic used in the final phase of the mediational analysis. The numerator is the product of the unstandardized regression coefficients for the paths for metamemory and Euclidean distance and between Euclidean distance and recognition accuracy ($\beta_s = 0.160 \times 0.074 = 0.0118$), and the denominator is the square root of $a^2 + b^2 + 2ab$ in which $a$ and $b$ are unstandardized regression coefficients and $ab$ is the corresponding standard error ($SE_{rep} = 0.0080$), $Z$ statistic ($Z = 1.62, p = .133$). (Use of Sobel’s technique for the calculation of mediation, in which the term $a^2 + b^2$ was omitted from the above equation, led to nearly identical results.) On the basis of these analyses, the indirect effect of metamemory on recognition accuracy is significantly reduced when Euclidean distance is entered as a mediator.

All subsequent analyses were repeated-measures logistic regressions with the general estimating equations: Metamemory accuracy predicted both recognition accuracy, parameter estimate was $1.544$, ($SE = .418$), Wald $\chi^2(1, N = 12) = 8.20, p < .001, p_{rep} = .99$, and Euclidean distance, $2.43 (p < .002, p_{rep} = .983)$. Further, Euclidean distance also predicted children’s recognition memory, Wald $\chi^2(1, N = 12) = 6.46, p < .01, p_{rep} = .99$. Finally, when Euclidean distance was added to the model, it continued to predict metamemory accuracy, Wald $\chi^2(1, N = 12) = 3.60, p < .05, p_{rep} = .97$, but it is important to note that the metacognitive predictiveness of recognition accuracy dramatically decreased, with its estimate declining from $1.544$ to $0.771 (p < .001)$. Thus, the predictiveness of metamemory accuracy was much smaller when Euclidean distance was taken into consideration ($F = 0.49, p = .30, p_{rep} = .638$), whereas Euclidean distance remained strongly predictive of children’s recognition accuracy ($p < .05; p_{rep} = .876$). As a mediator, Euclidean distance accounted for $52\%$ of the total effect of metacognition.

**Discussion**

The findings from this experiment are consistent with the hypothesis that representational constraints explain the age trends observed in the earlier experiments. These results provide specific evidence for the role of representational constraint on the effectiveness of metamemory, specifically episodic memory monitoring. Younger children were more accurate on metamemory judgments than were older children when the items to be judged were represented as farther apart for them than they were for older children. The reverse was true when items were represented closer to distractors for younger children than for older children, in which case the latter were far more accurate in their judgments. Substantively, this means that when the representation of distractors was close to the representation of critical items, both age groups were more likely to erroneously claim that the distractor was remembered and that the actual critical item was merely known or rejected. When the representation of distractors was farther apart from the representation for critical items, both age groups were far more accurate in metamemory judgments. In sum, younger children’s metamemory worked as well as older children’s when Euclidean distance was controlled. Thus, Euclidean proximity distance (an aspect of children’s representation) and not chronological age (a proxy for brain maturity) was associated with metamemory efficacy. It was Euclidean distance that mediated the role of metamemory in age differences in recognition accuracy.

It is important to note that in analyses that included age in the model, Euclidean distance was not only a strong predictor of metamemory accuracy that did not interact with age but rivaled age in its predictive potency, something rarely observed in the memory-development or individual-differences literatures (see Bruck & Melnyk, 2004). Thus, when a distractor item was represented for a given child as closer in semantic space to a critical item from the story, that child was more likely to erroneously claim that they remembered seeing the distractor. This was as true for older children as it was for younger ones. In fact, Euclidian distance was such a strong predictor that our a priori predictions of reverse-age effects in metamemory efficacy and recognition accuracy were nearly perfectly realized. When a distractor item was represented closer in Euclidian space to the paired critical item for older children then, contrary to usual age trends, older children were less able to monitor their memories accurately than were younger children. These predictions lend credence to the view that individual and developmental differences in underlying representations play an important role in metacognitive functioning (Orrstein & Elischberger, 2004).

Because representational status was a within-subject variable in the analyses, and it entirely mediated the effect of age, our data suggest that the level of neurocognitive maturation of the PFC, although perhaps important for more advanced forms of metamemory functioning than the monitoring task used here, is a secondary contributor to monitoring, exceeded by representational factors. These analyses indicate that the same child would behave differently as a function of the representational structures involved.

**General Discussion**

The present research was motivated by a counterintuitive pilot finding: Young children were sometimes unable to recognize a previously seen picture despite being able to correctly identify its attributes, such as where it was located. The awareness of a stimulus’ attributes is usually sufficient for adults to verify its prior occurrence and to distinguish it from nonoccurring stimuli (Johnson, 1988; Johnson, Foley, Suengas, & Raye, 1988). Barring some postevent suggestion, how could one be aware that a previously
presented item was located in the uppermost right-corner room in a two-dimensional house drawing and yet claim not to remember having seen the item itself. Adults who fail to recognize items despite recognizing their contextual attributes do so because of either encoding specificity mismatch (e.g., retrieval cues do not match one of the encoded meanings) or because of conservative decision criteria (e.g., setting threshold too high for a recognition hit; e.g., Muter, 1978; Starns et al., 2008). In contrast, the mechanism underlying preschoolers’ failure is not the result of encoding specificity, whereby different meanings of nominally identical items mismatch retrieval cues, because the same encoded meaning was involved here. Nor is preschoolers’ failure the result of using conservative decision criteria that lead adults to err (Starns’s et al., 2008). In fact, just the opposite was the case—for young children, the phenomenon occurs despite their amazingly high false-alarm rate (Experiments 1 and 2). If neither encoding specificity nor a conservative decision criterion was responsible for preschoolers’ failure to recognize pictures whose contextual attributes were recognizable then what was the mechanism? In addressing this question, we first make reference to various theoretical accounts of early memory, and in doing so, we contrast them with our own synthesis, the developmental–representational theory, which we offer as an answer to this question.

**Binding Theory**

There is growing evidence—both behavioral and neuroscientific—that an item’s attributes must be integrated with each other as well as with the item itself (Barsalou, 2008; Davachi et al., 2003; Lloyd et al., 2009). When cuing with an attribute is successful, this is evidence of attribute-item binding, and when an attribute fails to prompt the item, this could be the result of failure to encode the attributes in the first place (i.e., lack of binding the item and its attributes into a coherent memory record) or the disintegration of the bonds holding the attributes and item together during the retention interval (Howe, 2000). Sluzenski, Newcombe, and Kovacs (2006) provided evidence that preschoolers are particularly poor at binding together an item and its contextual attributes: 4-year-olds in their experiments could recognize items nearly as well as 6-year-olds, and they also could recognize the items’ contexts nearly as well. However, they had difficulty recognizing the items in their contexts—that is, they failed to bind the items with each other and, more to the point, they failed to bind the attributes of items with the items themselves. It is reasonable to assume that this developmental binding failure can result in the retrieval of attributes without recognizing the items to which they are bound. This fits the findings in Experiments 1 and 2, in which younger children were more likely than older children to recall attributes of unrecognized items, and suggests either that these attributes were not tightly bound together at the time of storage or that the binding disintegrated over time.

Our developmental–representational theory, which we describe below, shares with binding theory its developmental orientation (weaker bindings being less likely with increasing age). Also in agreement with binding theory, the developmental–representational theoretical synthesis posits that unbound attributes impede interitem memory. However, it does so via a specific claim: that the workings of metamemory render item retrieval based on stored contextual attributes more effortful and problematic for very young children. Our synthesis also diverges from binding theory in the presumed neurological role in attribute–item binding that has been proffered, particularly hippocampal maturation through age 5 years and PFC development over an even longer developmental period. As Sluzenski et al. (2006) argued,

> Supporting this hypothesis ( improvement in encoding and retrieving co-occurrence relations) is the literature on neural development, which converges on the claim that prefrontal cortical regions—areas known to be involved in both episodic encoding and retrieval (for discussions, see Wheeler, Stuss, & Tulving, 1995, 1997)—undergo slow development relative to other cortical regions. (p. 99)

(However, see Newcombe, Lloyd, & Ratliff, 2007 for a critical review of this neurological basis.) According to the developmental–representational theory, growth in these neural systems are conflated with age changes in representational development. When the latter is taken into account, children’s failure to bind is diminished, suggesting that sufficient maturation of these neural systems exist to permit successful encoding, monitoring, and retrieval.

**Trace Integrity Theory**

Trace integrity theory and its mathematical underpinnings have been invoked to assess independent processing during encoding, storage, and retrieval. As Courage and Howe (2004) noted, this literature collectively concludes that storage failure (rather than retrieval failure) dominates children’s forgetting, and it lessens with age as attributes are better integrated into a coherent memory trace. Our developmental–representational synthesis concurs with this theory in two important ways. First, it is consistent with the notion that attributes not tightly integrated with their items disintegrate over time, compromising storage and retrieval. Second, it offers a similar account of why older children’s traces are more integrated, namely, increasing knowledge. Unlike trace integrity theory, however, developmental–representational theoretical synthesis does not invoke neurological changes as the causal mechanism underlying storage maintenance and retrieval facilitation (see Courage & Howe, 2004, pp. 13–14, for description of neurological changes associated with trace integrity theory). Unlike trace integrity theory, this novel theory emphasizes the role of representational constraints on monitoring as the fundamental cause.

**Fuzzy Trace Theory (FTT)**

FTT posits that memory comprises a range of traces, from precise, literal, verbatim ones to gistlike or fuzzy ones. Research within FTT shows age-related changes in the relative weight of verbatim and gist memory, with verbatim memories being more common in younger children and gist memories being more common in older children and adults. FTT is one of the few other theories that has motivated reverse developmental effects (Brainard et al., 2008), in particular, in application to false memory. In FTT word learning tasks, older children are more likely to fall prey to semantic confusions between presented words and foils than younger children because the latter do not possess the relevant knowledge to understand these semantic connections. In contrast, the developmental–representational theory emphasizes failures to recognize previously presented items, despite being able to recog-
nize their attributes, and emphasizes the role of metamemorial failure, specifically monitoring failure, which results from limited representations on the part of young children rather than an abundance of knowledge on the part of older children. Another difference is that because the shift from verbatim to gist processing is assumed to occur during the elementary school years (Howe, 2000), FTT would have difficulty accounting for reverse developmental trends among preschool age children and among older children and adults. In contrast, the developmental–representational theory can account for reverse developmental trends as a function of age-cohort or individual differences in knowledge throughout the lifetime.

The Role of Neural Systems in Monitoring

In contrast to the theories described above, the developmental–representational theory emphasizes the constraint that knowledge places on the ability to monitor rather than on age differences in the ability to bind per se, the disintegration of bound attributes over time, or the constraints of immature PFC. This theory holds that item recognition failure occurs, despite awareness of item attributes, when children do not introspect into the workings of memory and monitor its contents. Although this is not contrary to the position taken by binding theorists, it stresses the role of monitoring failure rather than the role of binding failure. That is, it focuses on the situation in which memory’s ensemble of encoded contextual attributes is not integrated, thus preventing them from reaching awareness or activating each other. As noted, researchers have argued that the basis of metamemory deficits lies in neurodevelopmental immaturity in the PFC, which controls the assembling, tracking, and regulating of memories (e.g., Moscovitch, Kapur, Kohler, & Houle 1995). The PFC, as seen above, is also implicated by a number of alternative theories in episodic encoding and retrieval changes with age.

Developmental–representational theoretical synthesis takes issue with the claim that PFC immaturity is at the heart of metamemorial failure in these experiments. The PFC immaturity claim is supported by evidence showing that (a) monitoring aids memory, (b) the PFC is involved in monitoring (e.g., Hayama & Rugg, 2008, 2009), and (c) the PFC is still immature in 3-year-olds to 4-year-olds (Gogtay et al., 2004). Specific cognitive processes have different neural dependencies, with more basic cognitive processes depending on the earliest developing brain regions, and the higher level cognitive processes depending on later developing regions such as the PFC. However, this evidence is not sufficient to conclude that PFC immaturity is the source of young children’s monitoring and memory difficulties, just that it may be a source of such difficulties and not necessarily the primary source. The current studies identify another source: representational quality. Over the developmental period studied, it is a larger contributor to age-related variance in monitoring than is neurological maturation.

Three findings from the present study call into question the view that neurodevelopmental constraints play the primary role in age differences in monitoring. First, in Experiment 3, the usual developmental trend is flattened or even reversed when the representation of younger children is richer than that of older children. Second, within the same child there are some items for which they exhibit successful monitoring and others for which they do not, and the former occurs when younger children’s representations resemble older children’s (Experiment 4). Finally, when age differences in representational factors are controlled, the normal developmental trend is attenuated, and age does not account for variance in memory failure or monitoring.4

As noted, neuroscientific and cognitive memory researchers have argued for a three-step causal chain controlling memory functioning: (a) Monitoring aids memory, (b) the PFC is needed for monitoring, and (c) the PFC is immature in 3-year-olds to 4-year-olds. However, if an immature PFC is the source of young children’s monitoring and memory difficulties, then 3-year-olds and 4-year-olds should not outperform 6-year-olds (Experiment 3) and 9-year-olds (Experiment 4), at least on some items. Age should not lose its predictive power when representational richness is added to the model. The developmental–representational theory in Figure 4 differs from the traditional causal account in Figure 1, in which PFC functioning is the lynchpin. PFC $\Rightarrow$ metamemory $\Rightarrow$ recognition. Figure 4, although acknowledging the neural underpinnings of memory monitoring, emphasizes the role played by mental representations in how attributes are bound and brought into awareness during familiarity judgments.

In contrast to Figure 1, in which the advantages of item familiarity are dwarfed by those associated with neural maturation, here, the reverse is true. In support of this proposal, we presented evidence that very young children’s memory representations typically contain fewer contextual attributes and fewer multiple-attribute representations. It is clear from the present findings (by inference in Experiments 2 and 3, and explicitly in Experiment 4) that even the youngest children are able to use metamemory quite efficiently when their representations resemble those of older children, for example, by containing similar contextual attributes or being situated in similar Euclidean space. Not only did younger children have no metamemory deficit when their representations resembled older children’s but older children displayed metamemory deficits when their representations resembled younger children’s. The empirically compelling aspect of the present findings is the targeted reverse-developmental predictions that were confirmed in nearly all cases: Younger children outperformed older children when their representations were more elaborate and, hence, more congenial to monitoring. Moreover, the same child, regardless of age, was better at monitoring items whose representations were rich.

The results of the four experiments, taken together with the pilot study, are consistent with the theory that developmental changes in mental representations mediate the efficiency of recognition accuracy and metamemory efficacy. Developmental differences in factors other than representational ones (e.g., maturation of the various subregions of the PFC, undeveloped metamemory ability, poor strategy knowledge, increases in functional connectivity across subregions) cannot parsimoniously explain why age differences in monitoring and memory are dramatically reduced and, at times, even disappear when the mental representations of younger children resemble those of older children or why the same child’s

---

4 In these studies, age is a proxy for PFC maturity, given the extensive evidence that 3-year-olds and 4-year-olds have immature PFCs, and the physiological maturation of widespread PFC neural networks involved in working memory emerge well after age 4 years (e.g., Casey, Amso, & Davidson, 2006; Luciana & Nelson, 1998).
monitoring fluctuates according to the richness of her or his mental representation.

This does not mean that the regions of the PFC or optimization of functional connectivity play no role in the metamemory deficits. For example, they could underlie age-related memory and metamemory differences in children younger than the ones in the present studies, and/or they may influence more complex forms of metamemory than the monitoring studied here. Figure 4 illustrates that brain development does make monitoring and recognition somewhat more effective (see below)—but its contribution to explaining age-related variance is secondary to the contribution of representational richness. Thus, in the developmental–representational synthesis being proffered, neurodevelopmental maturity represents a necessary but not sufficient condition for effective processing, as we describe in Figure 4 below. However, to the extent that it is necessary in the present tasks, it appears to be early maturing, adding little advantage to that provided by representational factors.

In view of the well-known demonstrations of changes in brain morphology as a function of experience of London taxi drivers and jugglers, revealing larger hippocampal volume and increased gray matter density as their representations became more developed (Draganski et al., 2004; Maguire et al., 2000), one might take issue with our claim that PFC development explains only limited variance in metamemory or is sufficiently early maturing. Thus, in the developmental–representational synthesis being proffered, neurodevelopmental immaturity, it becomes necessary to have extremely rich representations to effectively engage in episodic monitoring (and in executive functions more generally), something unnecessary with greater neural maturity; (b) on a related point, it might be argued that the growth of functional connectivity with age allows older children to encode and retrieve richer representations; or (c) it might be argued that executive functions such as monitoring entail a set of component processes mediated by the various subregions of the PFC, each with its own developmental trajectory (Best, Miller, & Jones, 2009; Braver & Bongiolatti, 2002), resulting in unexpected nonlinearities with age—improvements over the preschool years, followed by decreasing performance during early adolescence and, occasionally, later decreases as well (e.g., Vakil, Blachstein, & Sheinman, 1998).

Such arguments are consistent with developmental neuroscience findings (frontostriatal connectivity increases with age and is correlated with executive task performance, e.g., Amso & Casey, 2006). However, they are unsatisfactory as an account of the present findings in the absence of additional assumptions to explain why, in the present study, older children, who typically monitor their memories effectively, behave like younger children when their representations resemble those of younger children’s. It is not obvious how an account centered on neurodevelopmental delay explains why 3-year-olds can effectively monitor and remember words like crayon, eagle, lobster, bear, horse, and robin but not words such as lemon, cow, cheese, deer, egg, and orange or why they can actually outperform older children when their representations are richer. Growth in functional connectivity over this age range cannot explain within-age variability, nor is it obviously a factor in between-children variability. This is because, notwithstanding the ample research documenting the involvement of the right dorsolateral PFC in monitoring (e.g., Hayama & Rugg, 2009) and the immaturity of this region in 3-year-olds and 4-year-olds (e.g., Luciana & Nelson, 1998), functional connectivity of the frontostriatal region does not parallel the age changes observed here. This region is still quite undeveloped in our oldest age group and does not mature until late adolescence or young adulthood (e.g., Amso & Casey, 2006; Best et al., 2009). If greater connectivity across neural regions constrains effective monitoring, it is unclear how the youngest children were able to effectively monitor rich representations. There is no evidence to suggest that the neural basis of monitoring is item specific or that the brain mechanisms involved in monitoring richly represented items mature earlier. Yet, absent such assumptions, it is not clear how the PFC-mediated claim would account for the present findings—hence, our sugges-
tion that neural growth is a secondary contributor to monitoring, with representational factors being the primary contributor.

Developmentalists have documented the importance of representations in early memory, for example, the onset of autobiographical memory influenced by the organization of knowledge (Howe, Courage, & Edison, 2003) or the offset of amnesia among 3-year-olds when their representations were flexible, so prior knowledge could be inferred in new contexts. This has led to a conclusion similar to the one we make here, namely that neural accounts of early forgetting may be unwarranted because the level of brain development needed for the task is low enough to be available even among young preschoolers whose knowledge is adequate (Hayne, Boniface, & Barr, 2000).

A Developmental–Representational Theory

The most parsimonious theoretical synthesis of the present findings with previous studies, including longitudinal pediatric structural MRI studies of the same children over time, posits that although neural systems involved in memory (i.e., the structures of the medial temporal lobe, PFC) undergo substantial development throughout childhood and well into adolescence and even young adulthood (e.g., Durston et al., 2006; Gogtay et al., 2004; Stevens et al., 2009), the levels needed for basic monitoring, binding, comparing, and organizing representations already exist by early childhood. Although further neural development of these structures undoubtedly influences more advanced aspects of cognitive and metacognitive development and may make monitoring somewhat more efficient, the present findings illustrate that it is not neural growth but the nature of mental representation that constrains the metamemory and recognition findings. This is true regardless of whether they are operationalized as Euclidean distance, richness of attributes, or flexibility in the deployment of knowledge structures (Hayne et al., 2000). Consistent with our interpretation, others have shown that signal changes in the right dorsal PFC region are related to accuracy monitoring (comparing target with nontarget strings); however, the correlation between monitoring accuracy and the signal change remained significant after controlling for age (Durston et al., 2006). This strongly suggests that factors other than chronological age or its neural correlates influence metamemory performance. These arguments lead to a set of interrelated hypotheses that form the core of a developmental–representational theory of early development.

First, knowledge representations exert an important influence on early cognitive development, including memory and metamemory, and possibly cognitive development writ large (reasoning, concept formation, theory of mind, rule integration). This is because representations constrain the efficiency with which processes can operate: Rich representations make relationships easier to detect, attributes easier to bind, items easier to chunk, and generalizations across domains apparent. Young children’s representations are typically impoverished due to their limited knowledge (Brainerd et al., 2008; Chi & Ceci, 1987), which translates into representations characterized by fewer attributes, which are less integrated with each other; there is less flexible deployment of them, and they are less integrated with other knowledge. However, sometimes preschoolers possess rich knowledge structures. Anderson (1991) described how preschoolers can reason in a more advanced manner than traditionally assumed if they possess rich representations of time, speed, and distance. When given information about the speed of a barking dog, the duration of its barking, and the distance it ran, 5-year-olds can reason algebraically, for example, Time = Distance/Speed, thus reasoning at a level assumed to be impossible according to Piaget’s theory. Developmental–representational theory holds that representational constraints on processing contribute primary and independent variance to memory and other cognitive tasks, compared with any variance contributed by neural maturation, which is secondary in the sense that it explains less of the age-related variance in monitoring and memory.

Second, because the prolonged period of neural development coincides with a prolonged period of knowledge acquisition needed to encode rich and flexible representations, this coincidence confutes the role that each plays in a given cognitive performance. This can lead to the overimputation of neural factors as opposed to representational constraints on otherwise adequate neural development. The proposed developmental–representational theory anticipates that when neural and representational factors are separated as a consequence of either experimental dissociations or naturally occurring reversals (e.g., when the stimuli such as Sesame Street characters are better known by younger children than older children—see Brainerd et al., 2008), chronological age, despite being highly correlated with changes in cortical and subcortical volume, will lose some of its predictiveness, and the role of representational constraints will be more evident.

Third, the lengthy rollout of neural development over the first 2 decades of life coincides not only with massive changes in knowledge acquisition but also with the unfolding of complex cognitive developments and information processing strategies. A number of neuroscience teams have now reported the results of repeated scanning of the same children over long periods of time, charting changes in activation patterns during performance of various cognitive tasks (see Durston et al., 2006). For example, Gogtay et al. (2004) scanned the same children’s brains repeatedly over an 8- to 10-year period and showed that the sequence in which the cortex matured mimicked the typical cognitive development milestones, “with the last to mature being areas involved in executive function, attention, and . . . (frontal lobes)” (Gogtay et al., 2004, p. 8178). During this rollout, many but not all neural systems steadily increase, with some, such as gray matter density in some regions, increasing through childhood and early adolescence then decreasing during the postpubertal years through synaptic pruning (Gogtay et al., 2004).

The proposed developmental–representational theory leads to the expectation that at any given point in cognitive neurodevelopment, there is an interplay between the regions of the brain that subserve the cognitive process in question and the representation on which it operates. Individual brain subregions follow temporally distinct maturation trajectories (Gogtay et al., 2004), rendering neurodevelopmental maturity of the brain region involved in a given process as a necessary but not sufficient precondition for effective processing. Figure 4 depicts this graphically. In the figure, processing efficiency—here the efficiency of processes involved in monitoring—are mediated by the richness of the representation. Even early developing brain regions, such as the ones involved in habituation and attention, already highly developed in infancy, will be inefficient at processing poor representations. It is proposed that later-developing systems, such as those involved in monitoring, are functioning in preschool children at
levels sufficient to monitor rich representations but not impoverished ones. For an effective cognitive performance, it is necessary not only to have the neural maturity needed for the cognitive process to function but to have encoded rich, flexible representations on which the process can operate. In the absence of such representations, performance will suffer. In Figure 4, the representational richness is reflected by the height of the ordinate, with higher indicating more attributes, greater dimensionality, and more integration across dimensions. Young children who are frequent viewers of Sesame Street represent its characters more richly than do older individuals who are not viewers, whereas the reverse is true of a domain such as weather, in which older individuals have a finer grained understanding.

Fourth, the developmental–representational synthesis anticipates reduced age differences and, at the most extreme, reverse age differences, when the nature of representations is similar across ages or favors younger children, provided that the relevant brain region for the process in question has matured to a sufficient level in all age groups. Along these lines, the theory anticipates sizable within-age and even within-child variability that is systematically linked to aspects of the items’ representations. For a given child, memory, metamemory (and other cognitive processes) will be better for items that are part of embellished representations than impoverished ones. Note that in such instances, there is no variance in brain maturity because it is the same child undertaking the same cognitive process, nor is there any variance in the availability of the process (e.g., monitoring ability) because, again, it is the same child who can exhibit the process in embellished representations but not in impoverished ones. The only variable in such situations is the representations associated with the items that were poorly, versus effectively, monitored.

Several subtleties in Figure 4 warrant mention. Returning to the neural maturation model in Figure 1, even rich representations of young children are placed low on the processing efficiency ordinate because, according to adherents of this model, preschoolers’ immature PFC should constrain monitoring. Thus, the line connecting preschoolers’ monitoring of Sesame Street representations to adolescents’ monitoring of weather representations has a pronounced upward slope, although there is a small advantage for processing richly represented items over poorly represented ones for all ages. Figure 4, in contrast, shows that although their respective rich representations and brain systems may be nearly the equivalent, adolescents will nevertheless monitor slightly better because of their use of other strategies, but only slightly. In Figure 1, the line connecting preschoolers’ impoverished weather to adolescents impoverished Sesame Street shows a steeply increasing slope because, even though adolescents do not monitor their impoverished representations as well as they do rich ones, the difference is dwarfed by the variance associated with PFC maturation; they monitor both rich and impoverished representations better than preschoolers monitor even their richest ones.

Although the model in Figure 4 is primarily driven by representational factors, the rich (top) and impoverished (bottom) examples for adolescents are both slightly more efficient than comparable examples for preschoolers. Even though representational characteristics determine most of the variance in monitoring, PFC is allowed in the model to contribute some variance—that is, adolescents monitor rich representations slightly better than preschoolers do, and adolescents also are slightly better at monitoring impoverished representations than are preschoolers, though neither age group is very effective at monitoring the latter. Hence, young and old children are efficient at monitoring and remembering rich representation but not impoverished ones; however, because of their more optimal synaptic density, the model allows for adolescents to be slightly more efficient when representations are equated across age. In rare situations when young children’s representations are richer than are older individuals, this will more than offset any neural disadvantage. The data from the current experiments are in closer accord with the representationally driven model than the PFC-driven model.

**Broader Theoretical Implications of a Developmental–Representational Theory**

The findings reported here strongly suggest neurological changes in the PFC are not the primary controlling factor in the growth of memory and metamemory performance over the period of early to middle childhood. The ability to monitor, regulate, and introspect into the contents of memory, which has repeatedly been associated with memory improvement, was shown to be mediated by aspects of the representation. If correct, this suggests a fundamental reconfiguration of assumptions about the course and nature of memory and metamemory development during early and middle childhood. Specifically, the lack of age-related differences observed in Experiments 2, 3, and 4 (when the representations were equated) challenges the traditional account of precisely what develops with memory development. As noted, reverse developmental effects are rare in the scientific literature and have never been reported for an aspect of metacognition.5

Ordinarily, younger children’s representations are impoverished compared with older children’s, containing fewer attributes (Ceci et al., 1980) and reduced dimensionality (Ceci et al., 2007). Classic theories posited that early development proceeds from syncretic and undifferentiated representations to highly differentiated ones (Werner, 1948; Werner & Kaplan, 1952). To a very young child, the world may be construed in terms of a few gross categories, such as size (large vs. small), evaluative valence (nice vs. naughty), animacy (living vs. inanimate), and so on. With development, these categories become refined, with myriad dimensions crosscutting them and serving as a basis for differentiation. For example, a food item can be differentiated in terms of its role (appetizer, entrée, side dish, dessert), dominant bimolecular structure (protein, fat, carbohydrate), and type (dairy, citrus, etc.). These categories not only differentiate related items but also serve as possible bases of integration and, as the present findings suggest, may lend themselves to easier mental patrol of their contents.

---

5 Reverse developmental effects have been demonstrated in memory performance (e.g., Ornstein & Elsichberger, 2004; Lindberg, 1980) when younger children’s knowledge of cartoon characters was richer than that of older children’s, leading them to remember more; when their interest in an object (e.g., sneakers) was greater, leading them to focus on it more (King & Yuille, 1987); or when a connected meaning paradigm made older children’s greater knowledge a liability (for review see Brainerd et al., 2008). However, they have never been demonstrated to occur with metamemory. In addition, they have not been shown to be mediated by the nature of children’s emerging representations.
Experiment 4 provided both qualitative and quantitative evidence that this classic syncretic-to-differentiated shift may characterize the developmental changes in children’s representations in the present study. Older children’s representations were more dimensionalized, with greater differentiation within a single category. For example, they possessed a predacity dimension that younger children’s representations lacked: In their triad ratings, older children sometimes did not group eagle with robin (or sparrow) if another predator was the third member of the triad (e.g., bear). In contrast, younger children always grouped eagle with robin (or sparrow), regardless of the third member of the triad (giving as their reason that they were both birds). Similarly, older children possessed more differentiated representations for foods, including a citrus dimension (lemon, grapefruit, orange), a dairy dimension (e.g., milk, eggs, cheese, butter), and so on, none of which was evident in younger children’s food representation (which included size, color, and liquid vs. solid). (This led younger children to group milk with soda instead of with other dairy items—eggs, butter, cheese.) Possessing only a few dimensions, younger children may suffer cue-overload effects when they attempt to resemble items in response to a small number of cues that do not narrow the search space (small things, red things, solid things, living things, nice things). This would be expected to result in closer proximity between a critical event and a distractor because they share the young child’s few and groser types of attributes. Consistent with this argument, when older children’s representations contained attributes that overlapped a distractor’s attributes, they were as prone to monitoring and memory errors as were their younger counterparts.

Our theoretical stance is that early cognitive development is characterized by the appearance of structures and processes, including metacognitive ones, which are initially tied to a single domain. With development, children begin the journey of noticing, inferring, and inducing the relevance of these domain-specific processes to other domains. For example, a 3-year-old’s halving rule may initially be tied to the domain of small, edible, round objects, prompting correct answers to questions such as, “If I cut an apple (orange, pear) in half, how many pieces will I have?” Such a rule may not be applied broadly, however (“If I cut a rug in half, how many pieces will I have?”), until the representation of halving is sufficiently differentiated to be applied across domains (e.g., work or energy), until they are integrated by crosscutting dimensions (Ceci, 1996).

This stance argues for the centrality of representational constraints in cognitive development. The present experiments extend this account to metacognition, a form of operative knowledge that is ordinarily viewed as transdomainal once it is acquired (Ceci, 1996). Experiment 4 revealed that the effectiveness of metamemomy is mediated by the representation on which it operates. It is important to note that representational development, like neural development, does not proceed evenly. Sometimes, younger children have richer and more dimensionality in a given domain than do older children, as in the case of cartoon characters that they watch regularly or preschool routines that they are exposed to daily. Under such conditions, we may anticipate improvements in metamemomy and memory, even to the point of occasional reverse developmental effects. However, such instances of younger superrior representations are the exceptions, and normally children acquire added dimensionality with age.

Thus, the role of representational constraints may have implications not only for recognition memory and metamemomy but also for the broader field of cognitive development, though the current program of research was focused only on the former. For example, age-related differences in concept formation, categorization, and reasoning may also be under the influence of representational differences between age groups. It seems reasonable to suppose that categorization and concept formation are influenced by the nature of the items’ representations, and there may be situations in which age effects can be reduced, akin to what was found here for memory and metamemomy. Failure to consider representational factors may result in the overimputation of various processes and schemes that are themselves mediated by aspects of the mental representations on which they operate.

Practical Implications

Although the theoretical implication of these findings is evident for memory development, there appear to be several practical implications that may not be evident. One is that children’s eye-witness testimony may be affected by the nature of their representations. Rather than assuming younger children are inevitably more suggestible than are older children and adults (for review see Ceci & Bruck, 1993), the present findings indicate that suggestibility may to some extent be situational, with younger children being less suggestible than older individuals when their representation of the original event and that of the suggested event are more distinct for them than they are for older persons (e.g., Ceci et al., 2007). A related implication of the present results concerns the credence attached to young children’s statements when they claim not to remember an event; the present findings indicate that they may have troves of information that lie below the radar of their mental patrol, which are potentially accessible via cueing.

Caveats

In this section we note four caveats: First, we have used the term representation quite broadly. Interitem Euclidean distance (Experiment 4) and number of attributes (Experiments 2 and 3) are limited aspects of an item’s representation, and other aspects may prove more important in future research. For example, at the behavioral level, the degree of integration and differentiation among attributes can also be viewed as aspects of a mental representation; at the neurophysiological level, the unique firing patterns that hippocampal neurons exhibit when remembering in a given context can be viewed as an aspect of a neural representation (Smith & Mizumori, 2006). The inability to take into consideration these other aspects of the representation is a limitation of the present research, but at the same time it makes it all the more remarkable that Euclidean proximity was as predictive as it was of developmental trends in recognition memory and metamemomy and was especially predictive of reverse developmental effects. Although the multidimensional scaling procedures served us well in Experiment 4, there are inherent limitations to their use to estimate semantic proximity between items. Euclidean distance is relative in multidimensional scaling procedures, always influenced by the other items being scaled. Representational space can change with the addition or deletion of even a few stimuli—a predacity or dairy dimension emerges only to the extent that the other items included evoke such dimensions. Thus, no confidence can be placed on the specific representations or age differences in them. Furthermore, representations are probably not static in their effect on memory but are dynamic and changing online (e.g., Dahan & Tanenhaus, 2004).
Second, metamemory measures correlate only moderately with each other, even among conceptually similar types of metamemory (e.g., mean $\gamma = .40$ for correlation between feeling-of-knowing scores and judgments-of-knowing scores), and it is likely that the construct itself may be multidimensional as some have argued (Leoncio & Nelson, 1990). To the extent that this is the case, different task variables can be expected to draw on different types of monitoring. Procedural differences in remember–know judgments can result in differences in the form of metamemory assessed (Eldridge, Sarfati & Knowlton, 2002), so the adapted format used here may assess a somewhat different aspect of metamemory than the traditional remember–know task, a compromise necessitated by the need to make the task suitable for young children. Again, to the extent that these limitations were at work in the present study, the findings may be an underestimate of the importance of representational constraints on age trends.

Third, there were many changes in procedure across these experiments that limit generalizations. Experiment 4 required children to remember details over an 8–10 day period, compared with the 2–3 day delay in Experiments 1–3, and it used story narratives rather than discrete pictures in houses. As noted, in Experiment 4 we used older children than those in the first three experiments. We also used a new operationalization of representation and a different memory measure than was used in the prior experiments. All of these divergences work against the proposed representational theoretical synthesis, as they should make it more difficult to observe reverse developmental trends: The wider age range in Experiment 4 renders the a priori predictions particularly daring because it was predicted that 4-year-olds would outperform 9-year-olds on half of the test items—an even larger age range than in the prior experiments. These changes also add generality to the proposed theory because representational richness was shown to operate across them. Lloyd and her colleagues (Lloyd et al., 2009), using a task considerably simpler than the ones used in this article, reported a set of robust findings that converge with those reported here. This is important not only from a theoretical perspective but also because it staves off the potential criticism that the findings reported here are paradigm-bound.

Finally, it is clear that items’ representations are not the only factor influencing developmental differences in recognition accuracy and metamemory. Strategies, declarative knowledge, processing speed, and other neurobiological changes may be instrumental as well and have been shown to unfold over the same age span studied in the present experiments (Luciana & Nelson, 1998; Pennington, 1997).

Future Directions

The locus of effects. The central question motivating these experiments was whether it is a lack of PFC development, a fundamental metamemory deficit, or an abundance of impoverished representations that is the primary constraint on preschoolers’ memory monitoring and recognition accuracy. The findings of Experiments 3 and 4, especially the meditational analyses, indicate that it is the latter and that even the youngest children had sufficient neural maturity to be able to effectively use metamemory to enhance their remember–know judgments and recognition accuracy when their memory representations were sufficiently rich. In view of the central aim of the present study, the experiments were not designed to answer questions regarding the locus of the metamemory-representation link—whether it is established during encoding (e.g., if younger children focused on fewer or different attributes than did older children) or during retrieval because an item’s attributes are not bound to each other and/or to the item itself. Although the present experiments were not designed to examine this question, there are hints that both loci may be involved in the metamemory-representation connection. Future research will be necessary to answer this question unequivocally.

Recent work by binding theorists has suggested that the locus of this effect occurs during retrieval when young children fail to bind the encoded contextual attributes to the item itself (e.g., Sluzenski et al., 2006). For example, Lloyd et al. (2009) found that improvement in memory binding between 4 years of age and 6 years of age was due to changes in retrieval as opposed to encoding. However, it is possible that the locus occurs prior to retrieval—during encoding or storage if younger children fail to focus on the item itself as much as on its attributes. As already noted, the current findings suggest that both loci may be involved. Consider that in Experiment 3, age differences were greatly reduced when the recognition test was restricted to items in which all three attributes had been encoded, suggesting that encoding sets the conditions for later retrieval efficacy. On the other hand, when only a single or double attribute was encoded, older children were significantly more effective at using these limited attributes to retrieve the item than were younger children, suggesting that retrieval efficiency makes a contribution that is independent of encoding. Taken together, these results are consistent with a dual-locus effect wherein younger children’s failure to monitor and recognize items whose attributes had been recognized was a consequence of both encoding fewer attributes (thus having fewer attributes to bind to each other and to the item itself) and having less attribute information available and/or less effective strategies at the time of retrieval. The monitoring requirements of the remember–know task can entail processes that occur at both encoding and retrieval and the right prefrontal region is associated with both episodic encoding and retrieval (Buckner, 1996).6

Manipulating representational quality. The neural underpinnings of regulatory behaviors such as monitoring are necessary but not sufficient. The default, when age differences in representational richness are not controlled, is to assume that younger children’s monitoring failure is the result of an inherent, neuro-
logically based inability to introspect into their memories. Future research could serve a crucial role by systematically taking representational richness into account when assessing such behaviors.

Barsalou (2008) has reviewed a large body of neuroscience research arguing that categorical knowledge is grounded in specific modal systems rather than in a single transmodal semantic memory. Knowledge about a robin’s or sparrow’s appearance is stored in a visual representation, its flight in a motor representation, its chirping in an auditory one, and so on. Thus, the representation of conceptual knowledge is viewed as a neural circuit in which shape, color, activity, and location are distributed across modalities, and the corresponding circuit that forms may become active in unison or only in part when one type of attribute is activated. This view of semantic memory is consistent with the theoretical position taken here, in that it implies that one attribute can be activated, although another may not be. It will be of great interest if future researchers evaluate the hypothesis that the organization of an item’s semantic representation determines whether it is activated independently from its color, location, and so on. Finally, virtually nothing is known about how monitoring tasks may be decomposed into their micro components (attention, detecting, initiating, regulating) and,—when they are recruited during task performance—can be tied to independent, structurally connected neural regions. This insight also has been voiced by others: “Almost any developmental change could be plausibly linked to changes in PFC, given that we do not know in a detailed way the neural circuit in which shape, color, activity, and location are distributed across modalities, and the corresponding circuit that forms may become active in unison or only in part when one type of attribute is activated. This view of semantic memory is consistent with the theoretical position taken here, in that it implies that one attribute can be activated, although another may not be. It will be of great interest if future researchers evaluate the hypothesis that the organization of an item’s semantic representation determines whether it is activated independently from its color, location, and so on. Finally, virtually nothing is known about how monitoring tasks may be decomposed into their micro components (attention, detecting, initiating, regulating) and,—when they are recruited during task performance—can be tied to independent, structurally connected neural regions. This insight also has been voiced by others: “Almost any developmental change could be plausibly linked to changes in PFC, given that we do not know in a detailed way the level of maturation in different areas needed to support specific functions” (Newcombe, Lloyd, & Ratliff, 2007, p. 18).

Conclusion

Until now, chronological age has been the single largest source of variance in developmental studies of memory and metamemory, with no other variable remotely approaching its influence (Bruck & Melnyk, 2004; Ceci & Bruck, 1993). The search for individual and group variables to account for developmental variance in memory has failed to find a source that rivals chronological age; younger preschoolers are less adept at a variety of memory and metamemory tasks, more suggestible, and have limited strategic use of attribute is activated independently from its color, location, and so on. Finally, virtually nothing is known about how monitoring tasks may be decomposed into their micro components (attention, detecting, initiating, regulating) and,—when they are recruited during task performance—can be tied to independent, structurally connected neural regions. This insight also has been voiced by others: “Almost any developmental change could be plausibly linked to changes in PFC, given that we do not know in a detailed way the level of maturation in different areas needed to support specific functions” (Newcombe, Lloyd, & Ratliff, 2007, p. 18).

References


Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children’s testimony: Can children control the accuracy of...
Appendix A

Regression Analyses

Because overall recognition accuracy was associated with an age-related difference in response bias, B, and P, it was important to take into account the extent to which such age differences in recognition accuracy were due to younger children using a more lax threshold for some cue levels than others and/or their discriminability being greater for some cue information than others. Toward this end, a composite measure was constructed to adjust for the level of cue that was associated with recognition errors (E). The findings from this analysis were fully consistent with the unadjusted analyses reported.

The composite simultaneously took into account each child’s correct recognitions, misses, and false alarms. Specifically, it included the percentage of each child’s correct choices, the level of cue information the child used to recognize each picture (room, quadrant and/or posterboard), and the percentage of false alarms to each level of cue information:

\[
E = \frac{\sum_{i=1}^{n} \beta r_{wi}}{N_{CR} + N_{M} + N_{FA}}
\]

where \(N_{CR}\), \(N_{M}\), and \(N_{FA}\) refer to the proportion of correct recognitions, misses and false alarms and \(\beta r_{wi}\) refers to a weighting factor that estimates the level of certainty a child requires before deciding that a picture is familiar, \(B_i = FA/(1 - (CR - FA))\). Lower values of \(B_i\) indicate more stringent decision criterion because failures to retrieve a picture in response to cues was more telling when the child’s level of certainty was generally high than when it was low, as might be expected.

This weight was determined by running a regression on pilot data we had collected using these same stimuli with a group of 17 children who were approximately the same age as those in Experiment 1. The dependent variable was a picture’s probability of being recalled in response to each level of cue, and the predictor variable was the cue level itself—room, quadrant, and posterboard. This analysis revealed that children of both ages were better able to recognize pictures if they could also correctly identify the room in which it was located, followed by the house’s quadrant in which it was located and, to a lesser degree, by the color of its posterboard. Thus, in the composite variable in Equation 2, the weighted error term (\(e_{wi}\)) is the difference between \(e_p\), errors in terms of levels of cues removed (1, 2, 3), and the probability of accurate recall to that level of cue, \(e_{Ai}\). Because proximal-level cues were more effective in prompting accurate recognition than were distal-level ones, level of cue information that led to accurate recognition was analyzed as a weighted variable, with each level weighted to reflect the unequal value of making an error at that level. The weights corresponded to the progression from the most proximal to the most distal levels of information. In this weighted analysis, the room in which the picture was located was relatively more proximal than the quadrant in which the house that contained the room was located, which in turn was relatively more proximal than the posterboard color that served as the backdrop. A regression was performed on each child’s recognition accuracy for each picture (\(n = 64\)) and its room (\(n = 4\)), house (\(n = 4\)), and posterboard color (\(n = 4\)), with these three levels of information as a predictor variable and with misses and false alarms excluded. The equation is

\[
e_{pi} = 0.22l + 1.45
\]

As expected, the level of information was a highly significant predictor of recognition accuracy, \(r(1012) = 98.4, p < .0001\), accounting for 22% of total variance in children’s recognition accuracy, and this did not differ reliably for the two age groups. Moreover, the
addition of a quadratic term failed to elevate this prediction (i.e., because there were only three levels of cuing, the quadratic is essentially the difference between the mean of the second level and the first level versus the difference between the mean of the second level and the third level). Both age groups were better able to recognize pictures if they could also correctly identify the room in which it was located, followed by the house’s quadrant in which it was located, and to a lesser degree, followed by the color of its posterboard. Thus, in the composite variable in Equation 1, the weighted error term ($e_{wi}$) was the difference between $e_{p}$ and the probability of correct recall to that level of cue. The observed age-related difference in response bias, $B_{r}$ and $P_{r}$ that was associated with recognition accuracy did not, therefore, alter the interpretation of the findings from this experiment.

Received April 22, 2009
Revision received December 21, 2009
Accepted December 22, 2009