



## PAPER

# Three-year-old children can access their own memory to guide responses on a visual matching task

Frances K. Balcomb and LouAnn Gerken

*Department of Psychology, University of Arizona, USA*

### Abstract

*Many models of learning rely on accessing internal knowledge states. Yet, although infants and young children are recognized to be proficient learners, the ability to act on metacognitive information is not thought to develop until early school years. In the experiments reported here, 3.5-year-olds demonstrated memory-monitoring skills by responding on a non-verbal task originally developed for non-human animals, in which they had to access their knowledge states. Children learned a set of paired associates, and were given the option to skip uncertain trials on a recognition memory test. Accuracy for accepted items was significantly higher than for skipped on a subsequent memory task that included all items. Additionally, children whose memory-monitoring assessments more closely matched actual memory performance showed superior overall learning, suggesting a correlation between memory-monitoring and memory itself. The results suggest that children may have implicit access to internal knowledge states at very young ages, providing an explanation for how they are able to guide learning, even as infants.*

### Introduction

Many current models of development rely on the implied assumption that infant behavior in learning paradigms reflects an infant's self-assessed level of competency with the information contained within a stimulus source (Houston-Price & Nakai, 2004; Hunter & Ames, 1988; Hunter, Ames & Koopman, 1983; Koenig & Echols, 2003; Koenig & Harris, 2005). Most of the research on the ability to access and act on information about one's own state of knowledge falls under the rubric of metacognition (Smith, Shields & Washburn, 2003; Washburn, Smith & Tagliabue, 2005). Children are not typically attributed with metacognitive abilities until late preschool at the earliest (Cultice, Somerville & Wellman, 1983; Flavell, Green & Flavell, 2000; Lockl & Schneider, 2002; Schneider, 1999), raising a potential paradox if accessing one's own knowledge is an important component of learning processes. A solution to this apparent paradox entertained here is that self-referential functions such as monitoring one's own memory are linked to core processes like memory, with the self-referential functions emerging implicitly in infancy along with the core processes. More specifically, we propose that metacognition may exist in infants as one undifferentiated system, and what is viewed as the emergence of metacognition at later ages may rather be the transition from a unitary system to a more well-differentiated, explicit system (Koriat, 1993, 1994; Nelson & Narens, 1990, 1994). We will explore existing evidence for this proposal in the

following sections leading up to three experiments, which provide evidence that children younger than previously thought (3.5-year-olds) can access information about how certain they are about the contents of their own memories. These results, in contrast to previous research findings (Cultice *et al.*, 1983; Flavell *et al.*, 2000; Lockl & Schneider, 2002; Schneider, 1999), suggest that young children can access their own knowledge states, albeit not explicitly, opening the possibility of metacognition in early development. This research, taken in combination with developmental research suggesting that infant behaviors in learning paradigms reflect some level of access to their own knowledge (Houston-Price & Nakai, 2004; Hunter & Ames, 1988; Hunter *et al.*, 1983; Koenig & Echols, 2003; Koenig & Harris, 2005), implies that memory-monitoring may indeed be present early in development, and may be a factor in early learning.

### *Uncertainty and learning in infants and children*

Infants and children are known to learn very quickly in their everyday environment and in complex experimental tasks such as language learning, in which there is no explicit goal, and little to no feedback guiding specific learning trajectories (Aslin, Saffran & Newport, 1998; Gerken, 2006; Gomez, 2006; Newport & Aslin, 2004; Thiessen & Saffran, 2003). Infants, like adults, appear to be more than mere associative learners and pattern matchers. They can select 'good' from 'bad' information sources (Campbell & Namy, 2003), and informationally

Address for correspondence: Frances K. Balcomb, Department of Psychology, University of Arizona, 1503 E. University Blvd, PO Box 210068, Tucson, Arizona 85721, USA; e-mail: francesb@u.arizona.edu

rich from coincidental co-occurrences as infants (Aslin *et al.*, 1998; Newport & Aslin, 2004). Indeed, infants' differential preference for what is novel vs. what is familiar in commonly used testing methods is typically attributed to their seeking an optimal level of new and old information to promote learning (Houston-Price & Nakai, 2004; Hunter & Ames, 1988). A standard way of interpreting direction of preference effects in infant research is that they reflect a tacit understanding by the infant of degree of learning, such that infants will show a novelty preference at test if they have mastered the information in training, and a familiarity preference if they have not. For example, in one study, 8- and 12-month-old infants were introduced to a complex array of toys, and allowed to play freely with them. One group was allowed to play with the toys until they were habituated (i.e. when their focused manipulation of objects dropped to a significantly low proportion per unit of time relative to prior unit of time) and compared to the second group, who were familiarized but not habituated to the toys. At test, infants were given both the familiar array and a novel array of toys to choose from. Infants who had habituated showed a significantly greater preference for novel toys than infants who had been familiarized but not habituated, suggesting that infants differentially sought out novel toys only after attaining sufficient experience with the familiar (Hunter *et al.*, 1983). This study suggests that infants have some control over how they are learning, using their level of competence to guide their behavioral responding to manipulate learning opportunities. In fact, direction of preference from familiarity to novelty can be reversed by increasing exposure to stimuli based on the assumption that increased exposure time will result in higher competency with training (Thiessen, Hill & Saffran, 2005).

#### *Models of adult metacognition*

To briefly summarize, traditional models of metacognition in adults, such as the Dual Process model, describe metacognition in adults as consisting of two separate components; a monitor and a control process, which act at an object level and a meta level (Nelson & Narens, 1990, 1994). Metacognitive monitoring provides information about the current state of cognitive processes occurring at the object level, resulting in such things as confidence judgments. Metacognitive control modifies activity at the meta level, for example allocating longer study times to certain items. Other models, such as the Trace Accessibility model, describe a relationship between memory and memory-monitoring such that metacognitive processes are derived from information that naturally arises during a memory search other than information about the availability of the target *per se* (Koriat, 1993, 1994).

Current theories of metacognition have expanded upon metacognitive models to describe optimization of learning under controlled study conditions. In a typical paradigm, people are given a set of items of varying

difficulty to learn under specified time constraints, and their study strategies are examined as a function of explicit metacognitive judgments. Theories resulting from this line of research describe a relationship between monitoring of item difficulty and learning rate, and control of cognitive study strategies. A person's study decisions are thought to be based on relative subjective difficulty of individual items and monitoring of one's learning to maintain an optimal rate (Dunlosky & Thiede, 1998; Metcalfe & Kornell, 2005; Son & Sethi, 2006; Thiede & Dunlosky, 1999), with a goal of optimizing learning by reducing maximal uncertainty. Thus learners allocate different strategies for items that differ in complexity, under varying time constraints. For example, the Discrepancy Reduction model proposes that people compare their level of actual competence to their level of desired competence, and focus their efforts on the hardest items, or those assessed to be at the furthest distance from the level of desired competence (Thiede & Dunlosky, 1999). The Region of Proximal Learning Theory (Metcalfe & Kornell, 2005) deals with learning in the intermediate area of difficulty, and states that learners self-monitor using two metacognitive processes, deciding which items to study and in what order, and deciding how much study time to allocate to each item. These current models of metacognition utilize explicit monitoring and control processes to self-guide learning and take advantage of the available structures in the environment.

Although conscious awareness is often an implied component of models of metacognition in learning, many researchers contend that consciousness is not necessary. Wellman (1977) introduced the idea of procedural metacognition, or implicitly acting upon self-knowledge, and indeed, there is evidence that monitoring processes may be implicit and may drive control processes (Diana & Reder, 2004; Reder & Schunn, 1996). Unconscious or implicit notions of metacognition have been highlighted in recent discussions relating metacognition, working memory, and executive function (Fernandez-Duque, Baird & Posner, 2000; Prins, Veenman & Elshout, 2006; Shimamura, 2000). The emphasis on implicit mechanisms of metacognition may provide an alternative approach to exploring uncertainty and self-monitoring processes, and the role these processes may play in learning early in life.

#### *Explicit metacognition in children*

Young children are not credited with strategic responding based on self-assessments until late preschool at the earliest, despite having developed an extensively organized knowledge base and cognitive strategies by then. However, much of the research on metacognition in development has been targeted at children old enough to demonstrate clearly explicit metacognitive processes through verbal report, and little research has focused on the availability of non-verbal metacognitive processes for guiding behavior at younger ages. The earliest evidence for explicit

metacognitive abilities in structured tasks with children does not emerge until age 4, and then only with use of familiar stimuli and task type (e.g. naming pictures of well-known, less well-known or unfamiliar children), or with indirect measures such as evidence of memory strategies like using semantic cues to aid memory (Cultice *et al.*, 1983; Sodian & Schneider, 1986). In more traditional metacognitive tasks, most research has been directed at school-aged children. This research reveals a strong developmental trend in metacognitive abilities, such that older children's assessments of what they know are more accurate than relatively younger children's throughout childhood (Lockl & Schneider, 2002). For example, in a feeling-of-knowing study, kindergarten, first, and third grade children were presented with line drawings, and asked to name them. If a child failed to correctly name an item s/he was asked if s/he would recognize the correct label if s/he heard it. Subsequently, children were presented with an array of nine pictures and one verbal label, and asked to point to the matching picture. Third grade children were significantly better at predicting which picture-label matches they would be able to recognize than first grade children, who in turn outperformed kindergarten children (Wellman, 1977). In general, young children tend to be overconfident in their metacognitive assessments, predicting a higher level of success than they actually achieve (Schneider & Pressley, 1997; Stipek & Mac Iver, 1989), confusing wishful thinking with realistic estimates of performance (Schneider, 1998), and confounding metacognitive estimates with actual memory contents (Butterfield, Nelson & Peck, 1988). The overall picture, when one focuses on explicitly measured metacognition, is one in which infants and young children do not appear to have access to their knowledge states, with emergence of rudimentary skills in early school years and developing until late childhood.

#### *Implicit metacognition in non-human animals*

The apparent lack of directed learning or metacognitive ability in human children is surprising, given that comparative research with non-human animals including dolphins, pigeons and monkeys provides evidence for implicit metacognition (Hampton, 2001; Inman & Shettleworth, 1999; Shields, Smith & Guttmanova, 2005). Many non-human animals demonstrate 'uncertainty' behaviors and can respond strategically, acting upon uncertain states to improve task performance. In one such study, Son and Kornell (2005; but see also Kornell, Son & Terrace, 2007) trained rhesus monkeys to use metacognitive monitoring judgments on a perceptual discrimination task. Monkeys were presented with a forced-choice discrimination, and immediately after the discrimination task indicated a retrospective confidence judgment about their discrimination response (high or low confidence). The high confidence option offered a gamble of a three token gain or loss, depending on whether the answer to the discrimination question was

correct, whereas the low confidence option offered them a one token gain regardless of accuracy. To test for transfer of metacognitive use, after monkeys were trained on the use of the metacognition confidence judgment task, they were presented with a novel perceptual judgment task, and also a working memory task. Monkeys were not only successful at using the confidence choice to indicate retrospective judgments of their own response accuracy, but were also able to transfer this skill to a task requiring a metacognitive judgment on a different underlying cognitive skill (working memory vs. perceptual judgment). In a subsequent task to test for metacognitive control, monkeys were trained in a sequence learning task in which they were presented with the option of taking a hint. Trials completed with no hints were rewarded with a high value food, and trials completed with a hint were rewarded with a lower value food. Results indicated that monkeys were able to use the hints optimally to improve their performance on trials that were more difficult, suggesting that they are capable of metacognitive control.

The results of metacognition in non-human animals support the idea that both metacognitive control and monitoring can be available as cognitive tools in implicit form. However, the metacognitive studies with children have typically relied on methodologies adapted from adult research, relying on explicit verbal measures. In contrast, the work with non-human animals has been specifically designed for non-verbal populations. Comparative research, revealing metacognition skills in monkeys that appear to outstrip those of young children, provides a clue as to a methodology for measuring metacognition in young children that may be more sensitive to their abilities. This suggests that the literature on metacognition in young children and infants may be more appropriately explored by utilizing paradigms designed to extricate implicit metacognitive processes, to determine whether children, like monkeys, can act strategically on uncertainty.

#### *Summary of existing literature*

The work with non-humans suggests that we need to re-examine our view that metacognitive skill exists primarily as a well-differentiated and explicit cognitive process that emerges late in human childhood and becomes superimposed on core cognitions. An alternative view proposed here, arising from the infant literature on learning and habituation (Houston-Price & Nakai, 2004; Hunter & Ames, 1988; Hunter *et al.*, 1983; Thiessen *et al.*, 2005) and the Trace Accessibility model of metacognition (Koriat, 1993, 1994), is that metacognition may be available as a cognitive tool for learning in the form of implicit access to knowledge states that can drive behavior, long before it is well differentiated and verbalizable. More specifically, what is hypothesized here is that early memory-monitoring, or the ability to access and utilize information about one's state of memory, may be a

resultant process of memory. On this view, some items are tagged as 'more certain' and others as 'less certain' during the retrieval process, based either on familiarity, a differential associated emotional valence, a sense of recognition, a threshold of recall strength, or another factor. Thus infants' control over cognitive behaviors, such as sustaining attentional focus (Houston-Price & Nakai, 2004; Hunter & Ames, 1988), may be influenced by implicit access to internal states of knowledge, with the goal of most efficiently maximizing learning while avoiding redundancy. If, indeed, metacognitive monitoring and control can be subserved by a simple mechanism integrated into the memory process itself, we might begin to understand how infants and young children appear able to capitalize on information in memory to learn rapidly and efficiently, without demonstrating explicit awareness of their states of knowledge. If children can be shown to similarly access internal knowledge states to guide behavior at an age when they typically fail explicit metacognitive tasks, we will have evidence that such higher level executive function skills may be available as tools for self-guided learning much younger than would be indicated by explicit measures. If children are able to use this information to strategically direct learning in optimal trajectories, we could begin to explain their very rapid learning at very young ages. The purpose of the current research was to examine the implicit, or non-verbal, relationship between memory and memory-monitoring in preschool children using methods adapted from comparative literature.

## Experiment 1

The purpose of Experiment 1 was to examine whether children are able to indicate evidence of implicit, or non-verbal, memory-monitoring in a recognition memory task by selectively accepting or declining trials based on memory of paired associates. The methods for this study were directly adapted from a memory-monitoring paradigm with rhesus monkeys (Shields, 1999; in Smith *et al.*, 2003). In this study, children learned a set of paired picture associates. During the memory-monitoring test, they were shown one item of a pair and chose to either accept or decline taking a trial in which they selected the mate of that item. Memory of paired associates was examined in a subsequent memory test in which children had to select one of two objects (the mate vs. a foil) on every trial. Each foil object was a mate for a different paired associate. Since both mate and foil were equally familiar and viable candidates for at least one pair, this method ensured that children were not able to select the mate on a given trial based on familiarity, or on knowledge of which objects were more likely to be targets. It was predicted that children would show poorer memory for the trials that they had declined in the memory-monitoring task than for trials they had accepted. In addition, Experiment 1 explored the possible relation between

memory-monitoring and memory by asking whether those children who made the most accurate assessments of their memory also showed better memory for the visual pairs than children who either overestimated or underestimated their knowledge of the pairs.

### Participants

Participants were 25 children aged 3:5–3:7 ( $M = 3:6$ ).

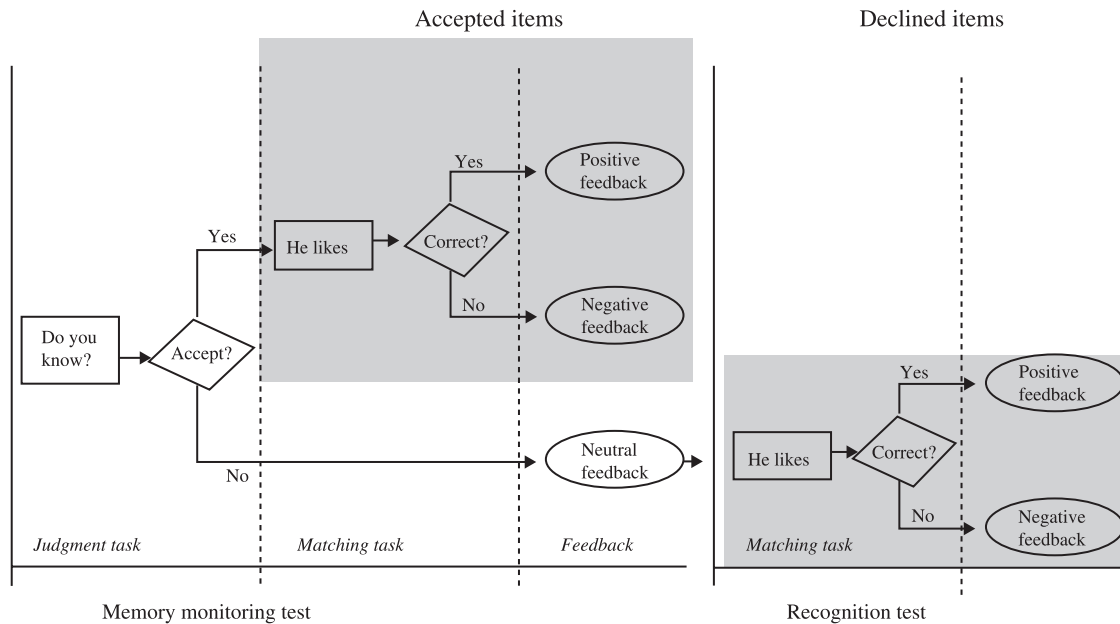
### Materials

The material to be remembered consisted of 15 visual pairs presented in the form of a movie. A pair consisted of a colored line drawing of a novel animal paired with a picture of a common object like a bicycle. Audio files for training and test stimuli were recorded by a female native English speaker using Sound Studio 2.1.1. Pairs were presented in movie format created with iMovie HD on a Macintosh Mini computer. First the picture of the animal appeared with a sentence giving its name (e.g. 'This is Andy'). Next, an object appeared with a sentence describing the animal–object relationship (e.g. 'He likes to drive a fire truck'). Finally, the animal and object appeared together with an additional sentence about the relationship (e.g. 'He drives a fire truck to all the fires'). Each picture appeared for 4 seconds, with a 1 second gap between.

Testing was conducted on a Dell Inspiron 1100 laptop with a Keytec Magic Touch touch-screen mounted on a 17" LCD upright monitor. The experiment was run using the DMASTR software developed at Monash University and at the University of Arizona by K. Forster and J. Forster. Test materials consisted of seven training trials, which included three animal–object pairs of familiar real animals and objects (e.g. a dog and a bone) and four of the 15 animal–object pairs from the movie. Actual test trials used the remaining 11 animal–object pairs. Test stimuli consisted of the animal–object pairs, and also included a line drawing of a colored arrow. The arrow served as an 'opt out' option, such that children could touch the arrow to skip trials. Audio stimuli for test consisted of the questions, 'Do you know what s/he likes?' and four types of positive (e.g. 'yay!'), one negative ('uh-oh, that's not it!') and one neutral ('we don't know what s/he likes') feedback sound file.

### Procedures

Children watched the movie once when they came to the laboratory. Subsequently, they went into a different room for testing. There were two separate tests; the first test was a *memory-monitoring test* consisting of a judgment task followed by a recognition memory (matching) task for those trials that the child accepted. The second test was a separate recognition memory test (*recognition test*), which assessed children's memory for all 11 of the test items (see Figure 1).



**Figure 1** Test procedures for Experiments 1 and 2. Although children took all trials on the recognition task, data for accepted items were taken from the memory-monitoring test, and data for declined items were taken from the recognition test (indicated in grey).

On the *memory-monitoring test*, children had two potential tasks. The judgment of memory (*judgment task*) was performed on every trial. Children were shown a picture of one of the animals from the learning phase, an arrow, the object mate of the animal, and a foil object (which was an associate for a different animal) and were asked 'Do you know what s/he likes?' Immediately after the question, the mate and foil disappeared, and children selected the animal or arrow with only those two images present. They touched the arrow on the touch screen if they thought they did not remember the item that had been paired with the animal (indicating a 'decline' response), which resulted in a new trial. They touched the animal if they thought they remembered the paired item (indicating 'accept'), which started the second task, the *matching task*. Note that children took the matching task on the memory-monitoring test only if they indicated an 'accept' response. On the matching task, the arrow disappeared and children selected the correct match from the target (mate) and foil object. If a child chose the correct response (the mate), s/he received positive feedback, but if s/he selected the incorrect target (the foil), s/he received negative feedback.

After the memory-monitoring test, all children were presented with an 11-trial recognition memory test (*recognition test*) in which they saw the same animals as on the memory-monitoring test. On each trial, children saw one of the 11 animals, the animal's object mate and another object that was the mate of a different animal. They were simply asked to select which object had been paired with the animal. For some children, some trials of the recognition test involved being tested on an animal

a second time if they had accepted the trial involving that animal on the memory-monitoring test.

There were two training test sets to ensure that children understood the procedure. The first training set consisted of three common animals and common associates (e.g. a dog and a bone, with eyeglasses as a foil), and three novel imaginary animals. The experimenter completed the first training set with children, expressing knowledge of and ensuring choice of the familiar animals, and expressing ignorance of and ensuring decline of the novel imaginary animals. To encourage children to transfer the procedures to the animals from the movie, the second training consisted of four animals; the first and last animals of the movie (which, according to list learning effects should be better accessible to memory) and two animals from the middle of the movie, which should similarly be less accessible. These items were not used in the subsequent tests. Children completed the second training set with the experimenter verbally reviewing the child's responses after each item. The actual tests (memory-monitoring and recognition) consisted of 11 pairs of animals and objects, which children completed without help. After each trial, the experimenter echoed the feedback given by the computer.

### Results

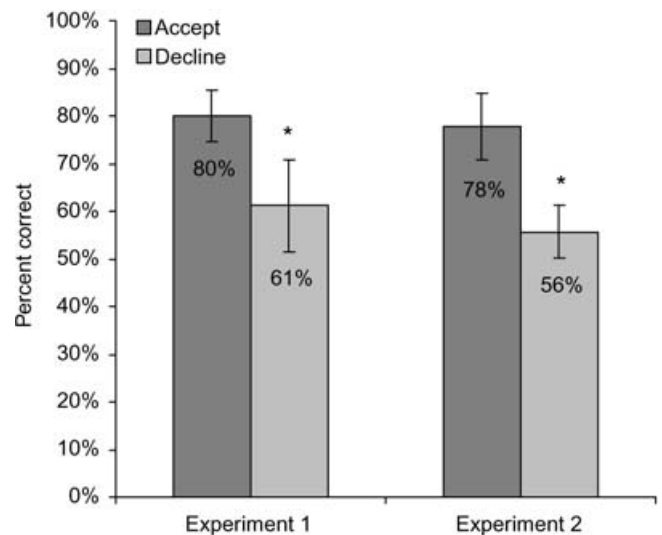
All children took both the memory-monitoring and the recognition tests. The response (correct or incorrect) from the first time a child encountered the matching task for each trial was used, regardless of whether it occurred on the memory-monitoring or the recognition test

(indicated in grey on Figure 1). For accepted trials, responses came from the matching task on the memory-monitoring test. For declined trials, responses came from the later recognition test. These data were used for all analyses presented here.

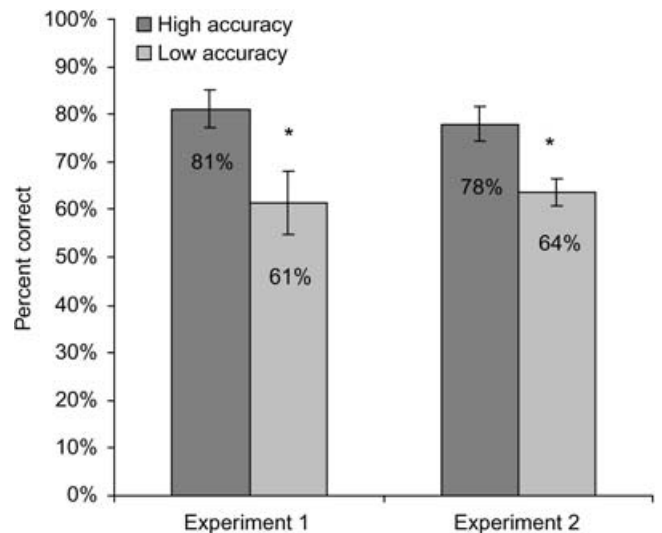
One child declined all trials; thus his data came from the recognition test only. Eight children accepted all trials, and thus their data came from the memory-monitoring test only. Sixteen children accepted and declined at least one trial, and their data thus were combined from both tests. Using the child's first response, regardless of which test it came from, was done in order to avoid learning that might occur between the memory-monitoring and recognition memory tests for items that children accepted on the memory-monitoring test. Because the same 11 items were used as both matches and foils, it is possible that learning during the memory-monitoring test also improved performance on later trials and on the recognition test for previously declined items, simply by process of elimination. However, the potential for such learning on declined items works against our hypotheses. Therefore, we deemed this approach the most conservative one.

Children's mean accept rate was 8.1/11 trials (74%), and mean overall accuracy was 7.9/11 trials (72%). The first analysis was conducted to ensure that performance was significantly above chance. Overall accuracy was compared to chance performance, revealing that children were significantly more accurate ( $M = 72\%$ ,  $SD = 21$ ) than chance ( $M = 50$ ,  $t(48) = 5.14$ ,  $p < .001$ , two-tailed). An item analysis revealed that accuracy on accepted items ( $M = 73\%$ ,  $SD = 19$ ) was significantly higher than chance ( $M = 50\%$ ,  $t(10) = 3.96$ ,  $p < .01$ , two-tailed). The next analysis was conducted to determine if children demonstrated better accuracy at selecting the correct mate for items they accepted than on those they declined on the memory-monitoring test, indicating accurate judgment of their knowledge. For the 16 children who both accepted and declined at least one trial, the percent correct on the matching task of the memory-monitoring test for accepted trials was compared to the percent correct on the recognition test for items children declined on the memory-monitoring test. As predicted, children were significantly more accurate on 'accept' items ( $M = 80\%$ ,  $SD = 22$ ) than 'decline' ( $M = 61\%$ ,  $SD = 39$ ;  $t(15) = 2.12$ ,  $p = .05$ , two-tailed,  $d = .53$ ; see Figure 2).

The next analysis was conducted to examine whether better memory-monitoring performance was related to better memory performance. Each child selected a mate for all 11 animals across both tests; therefore the total number of correct responses was known. For each child, the number correct on the 11 trials for which a child selected a given animal's mate for the first time (again using data from both the matching task of the memory-monitoring test for accepted items, and the recognition test for declined items) was compared to the number of items the child expected to get correct, as indicated by the number of items s/he accepted on the memory-monitoring test, to determine overall metamemory performance. A



**Figure 2** Percent correct for accept type trials vs. decline on Experiments 1 and 2.



**Figure 3** Percent correct on all trials for high accuracy vs. low accuracy groups in Experiments 1 and 2.

difference score was calculated (total correct minus total accepted) and the absolute value taken to determine distance from zero. Children were ranked as to how far their performance estimates differed from their actual performance, with zero (no difference) being optimal. Children were divided at the median into two groups, 'higher accuracy' (HA) and 'lower accuracy' (LA), and the number of correct mates selected by HA versus LA children was compared. Equal numbers of children who accepted all items fell into each group, and thus any effect of accepting high numbers of trials was distributed. Results indicate that children in the HA group selected the correct mate significantly more often ( $M = 81\%$ ,  $SD = 15$ ) than children in the LA group ( $M = 61\%$ ,  $SD = 23$ ;  $t(23) = 2.61$ ,  $p = .02$ , two-tailed,  $d = 1.01$ ; see Figure 3). Consistent with the  $t$ -test, a Spearman rank order correlation

between the absolute value of the difference score and the total percent correct revealed a significant relationship,  $r(24) = .45$ ,  $p \leq .02$ . These results suggest that children who either over- or underestimated their memory abilities showed poorer overall memory, although the possibility that children who are better at estimating their memory demonstrate a combination of better learning and memory cannot be ruled out.

## Experiment 2

The results of Experiment 1 indicate that children are able to evidence memory-monitoring in a recognition memory task by selectively accepting or declining matching trials based on memory for matched pairs. This design was modeled after the task with which rhesus monkeys were successful at indicating metacognitive competence on visual matching tasks (Shields, 1999; in Smith *et al.*, 2003). However, in that study, monkeys were able to decline less well-known items and accept more well-known items only when they saw the target, the match and the foil at the time of choice, but not when they chose to accept or decline trials in the absence of the target and foil. The purpose of Experiment 2 was to further explore the relationship between memory-monitoring and memory by examining whether children can show evidence of memory-monitoring in a retrieval memory task, in which monkeys failed. This study was identical to Experiment 1 except that on the memory-monitoring test, children chose to accept or decline trials with only the animal and arrow present, without seeing the target and foil. One reason to believe that children might succeed where rhesus monkeys failed is that in both metacognitive tasks, monkeys required many hours of training and thousands of test trials, whereas children required two short trainings and only 11 test trials. Further, young children have access to language and the benefits of a symbolic representation system, perhaps providing them with a cognitive advantage over monkeys (Smith, Minda & Washburn, 2004). Therefore, children may be completing the tasks differently from monkeys.

### Participants

Participants were 29 children aged 3:5–3:7 ( $M = 3:6$ ). Data from three children were discarded due to inability to complete the task unaided.

### Materials

The materials were the same materials used in Experiment 1. The memory-monitoring test differed from Experiment 1 in that the judgment task was completed without viewing the mate and foil, but with only the animal and arrow present. There were two training sessions followed by the test. In the first item of the first training children were asked, ‘Do you know what he likes

– do you think you could guess if you saw the matches?’ As in Experiment 1, the experimenter completed the first training with children, expressing knowledge of and ensuring choice of the familiar animals, and expressing ignorance of and ensuring decline of the novel imaginary animals.

### Procedures

The procedures were the same as Experiment 1.

### Results

Two children declined all trials, 11 children accepted all trials, and 16 children accepted and declined at least one trial. Children’s mean accept rate was 6.6/11 trials (60%), and mean overall accuracy was 7.7/11 trials (70%). The first analysis was conducted to ensure that performance was significantly above chance. Overall accuracy was compared to chance performance, revealing that children were significantly more accurate ( $M = 70\%$ ,  $SD = 15$ ) than chance ( $M = 50\%$ ,  $t(56) = 7.40$ ,  $p < .001$ , two-tailed). An item analysis revealed that accuracy on accepted items ( $M = 77\%$ ,  $SD = 24$ ) was significantly higher than chance ( $M = 50\%$ ,  $t(10) = 3.83$ ,  $p < .01$ , two-tailed).

The second analysis was conducted to determine if children demonstrated better accuracy on items they accepted than on those they declined on the memory-monitoring test, indicating accurate judgment of their knowledge. For the 16 children who both accepted and declined at least one trial, the percent correct on the matching task of the memory-monitoring test for accepted trials was compared to the percent correct on the recognition test for items children had declined on the memory-monitoring test. As predicted, children were significantly more accurate on ‘accept’ items ( $M = 78\%$ ,  $SD = 22$ ) than ‘decline’ ( $M = 56\%$ ,  $SD = 22$ ;  $t(15) = 2.7$ ,  $p = .02$ , two-tailed,  $d = .68$ ; see Figure 2).

The next analysis was conducted to examine whether better memory-monitoring performance was related to better memory performance. As in Experiment 1, absolute values of total correct minus total accepted difference scores were computed for each child, and children were divided at the median into two groups, ‘higher accuracy’ (HA) and ‘lower accuracy’ (LA). Seven children who accepted all trials fell into the HA group, and four were in the LA group. The overall accuracy of HA versus LA children was compared. Results indicate that children in the HA group demonstrated significantly higher overall accuracy scores ( $M = 78\%$   $SD = 14$ ) than children in the LA group ( $M = 64\%$ ,  $SD = 11$ ;  $t(26) = 3.10$ ,  $p = .005$ , two-tailed,  $d = 1.11$ ; see Figure 3). Consistent with the  $t$ -test, a Spearman rank order correlation between the absolute value of the difference score and the total percent correct revealed a trend in the expected direction,  $r(27) = .36$ ,  $p < .06$ . As in Experiment 1, the data suggest that children who either over- or underestimated their memory abilities showed poorer overall memory skills.



### Experiment 3

The results of Experiments 1 and 2 indicate that children are able to evidence memory-monitoring in a recognition memory task by selectively accepting or declining matching trials based on memory for paired associates. Despite their more conservative acceptance rates when the task involved recall (Experiment 2) as opposed to recognition (Experiment 1), they showed a significant ability to monitor their own memory in both tasks. However, in both experiments, children's accuracy for 'accept' items was measured using data from the memory-monitoring test, whereas accuracy for the 'decline' items was measured using data from the subsequent recognition test. This situation raises the possibility that memory performance on items that were declined may have been equivalent to that of accepted items when the accept/decline choice was originally made, and subsequently worsened due to forgetting or interference during testing, thus artificially lowering the later observed accuracy for 'decline' items. Experiment 3 was designed to measure the memory performance of children for paired associates immediately after viewing the movie. If children's accuracy on declined items was lowered as a function of interference, forgetting, or receiving differential feedback on those items during the memory-monitoring test, then accuracy for the weighted average of all items in Experiment 2 ('accept' and 'decline') should be significantly lower than accuracy on a memory test taken immediately after viewing the movie, in which no items have had the opportunity to worsen in memory and thus pull the overall average down. If, however, consistent with our prediction, the poorer performance on declined trials at the time of the recognition test reflected poorer knowledge of those items at the time the choice was made on the memory-monitoring test, then overall memory accuracy on the memory-monitoring test in Experiment 2 ('accept' and 'decline') should not differ from performance of children who take only the recognition test. Further, memory for items accepted during the memory-monitoring test of Experiment 2 should be higher than overall memory for children who are tested immediately after the movie, since the latter test includes items that children presumably would decline if given the choice.

#### Participants

Participants were 27 children aged 3:5–3:7 ( $M = 3:6$ ). Data from two children were discarded due to inability to complete the task unaided.

#### Materials

The materials were the same as used in Experiments 1 and 2, with the exception that on the test phase only the recognition test was utilized. The memory test was introduced by telling the children that they were going to play

a matching game with the animals from the movie they had just watched.

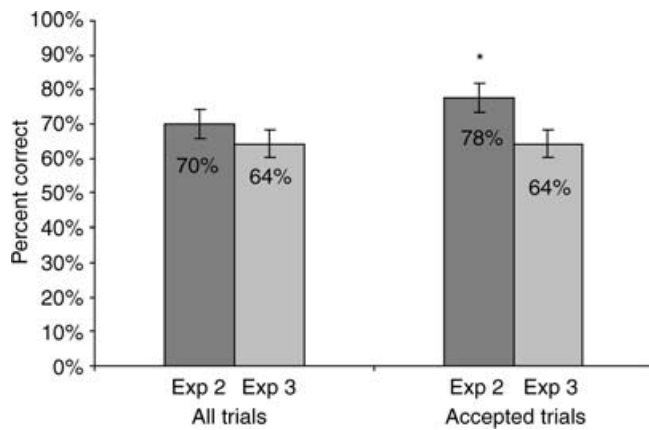
#### Procedures

The experiment consisted of a learning phase and a test phase. In the learning phase, as in Experiments 1 and 2, children watched the 15-item movie once. In the test phase, children took a recognition memory test in which they saw the same 11 items that had appeared on the memory-monitoring test in Experiments 1 and 2, to assess their memory for the paired associates seen in the learning phase. The experimenter explained the game to the children by saying, 'We're going to play a matching game. We're going to see some of the animals from the movie, and I want you to find the match for each animal.' Children were shown a picture of one of the animals from the learning phase, the object mate of the animal, and a foil object (which was an associate for a different animal) and heard the prompt, 'He likes . . . ?' For the first trial only, after the auditory prompt the experimenter encouraged children to find the mate by asking, 'Do you know what he likes?' The experimenter then pointed to both the target and foil, asking 'This, or this?' No further prompting was given, although, as in Experiments 1 and 2, the experimenter echoed the feedback provided by the computer after each trial.

#### Results

The first analysis was conducted to compare children's memory accuracy in Experiment 3 with their accuracy from Experiment 2, which combined accepted items from the memory-monitoring test and declined items from the memory test. No difference was predicted. Data from all 29 children from Experiment 2 were used. The weighted mean accuracy for each child in Experiment 2 was calculated by using data from the memory-monitoring test for items that were accepted during that test and data from the recognition test for items that were declined during the memory-monitoring test. For children who accepted all trials, the data came from the memory-monitoring test only, and for children who declined all trials, the data came from the recognition test only. The total percent correct for Experiment 2 was compared to the percent correct for Experiment 3. Counter to the view that children's weaker performance for declined items in Experiment 2 was due to forgetting or interference, the percent correct in Experiment 3 ( $M = 64\%$ ,  $SD = 21$ ) was actually lower than that in Experiment 2 ( $M = 70\%$ ,  $SD = 15$ ), although not significantly so ( $t(54) = 1.18$ ,  $p = .24$ , two-tailed,  $d = .19$ ; see Figure 4). This result suggests that the memory performance of both groups reflects the same level of ability, and provides evidence against the view that children in Experiment 2 had worse performance on declined items by the time they were tested in the memory test.





**Figure 4** Percent correct for all trials in Experiment 2 vs. Experiment 3, and accept type trials in Experiment 2 vs. Experiment 3. Note that in both comparisons all trials on Experiment 3 were used, since children were forced to take all trials.

Children in Experiment 2 were able to accept and decline trials for which they selected the paired associate for the animal. Twenty-seven children accepted at least one trial. In the original analysis of Experiment 2, the data for the 11 children who accepted all trials could not be included in the within-subjects analysis comparing performance on 'accept' vs. 'decline' trials. However, a between-subjects analysis could be used to compare the memory performance of children in Experiment 3 to the performance on 'accept' trials of the 27 children who accepted at least one trial in Experiment 2. If children selectively accepted trials in Experiment 2 based on memory-monitoring, accuracy on accepted trials should be higher than the accuracy on the forced-choice memory test from Experiment 3. The latter test presumably included items that children would have declined, as well as items they would have accepted, if given the choice. As predicted, children were significantly more accurate on 'accept' items from Experiment 2 ( $M = 78\%$ ,  $SD = 22$ ) than on the forced-choice memory test of Experiment 3 ( $M = 64\%$ ,  $SD = 15$ ;  $t(52) = 2.8$ ,  $p = .03$ , two-tailed,  $d = .74$ ; see Figure 4).

Taken together, the two analyses comparing the results of Experiment 2 and Experiment 3 suggest that, when children are allowed to selectively choose items to accept and decline on a memory-monitoring task, they are able to effectively use the opt-out option to improve their performance over that of children who are not given a similar opt-out option.

## Discussion

These experiments indicate that young children are able to demonstrate evidence of implicit memory-monitoring skills before the age at which they can verbalize about their knowledge. In addition, the results suggest that memory-monitoring may be linked to memory itself.

The results are a first step toward addressing the paradox of excellent learning abilities in infants and young children despite previous findings demonstrating poor metacognitive skills (Cultice *et al.*, 1983; Flavell *et al.*, 2000; Lockl & Schneider, 2002; Schneider, 1999). The picture that is beginning to emerge is one in which learners are able to keep track implicitly of what they do and do not know from very early in life. Some aspects of this ability can be tapped in non-verbal metacognition tasks such as ours by the preschool years (and perhaps before). However, the ability to verbalize about the contents of one's memory develops later. Thus, knowing and being able to make use of the contents of one's mind represents a continuum of task-particular abilities.

Although the current study only demonstrates that children possess the ability to access information about their memory earlier than has previously been proposed, these results do not explain how children accomplish the observed memory-monitoring abilities. One possibility is that perhaps, in generating a memory search, children used different markers than a feeling of knowing *per se* as decision criteria. For example, children may have chosen to accept trials based on a sense of 'liking' for certain items over others. An alternative explanation is a simple threshold account. By this account, when a child generates a memory search that results in an activation of a memory representation above a certain threshold the child experiences successful recall and chooses to 'accept'; otherwise s/he experiences recall failure and responds 'decline.' However, perhaps it is these very 'alternative' explanations that comprise memory-monitoring. In fact, in support of this hypothesis, there is evidence that repeated exposure to a stimulus can result in a change in affect toward that stimulus, known as the exposure effect (Zajonc, 1968). Perhaps it is this familiarity/liking effect that potentially provides one way of differentiating known from unknown items, or items differing in levels of uncertainty. By this account, a feeling of knowing can arise from a feeling of liking, or a feeling of liking may be what marks an item as being 'known', and disliking what marks it as 'not-known'. What is interesting is that these mechanisms appear to be successful, as evidenced by the accuracy on accepted versus declined items.

A second finding from this research was that, in both Experiments 1 and 2, children who demonstrated better memory-monitoring performance also demonstrated better overall recognition memory performance than children who either over- or underestimated their memory. This intriguing finding suggests that metacognition is linked to the process on which it acts, such that performance in one domain is inherently linked to performance in the other. This linkage further suggests that individuals who differ in core processes like memory may also differ in the control and monitoring processes that act upon them. In support of this hypothesis, recent focus on individual differences reveals that older children and adults also vary in their ability to optimally respond to uncertain

tasks, and that this difference may be correlated with other cognitive and personality variables (Schneider, Kron, Hunnerkopf & Krajewski, 2004; Sodian & Schneider, 1986; Washburn *et al.*, 2005). The current results, taken together with extant literature, suggest that what has typically been called 'metacognition' and cognition are inherently interrelated. Further, the results of this experiment open the possibility that metacognition may emerge implicitly very early, in fact as early as when core processes emerge, providing a mechanism by which infants and children can actively direct learning.

It is unclear what underlies the individual differences resulting in the disparity in performance observed in these experiments. It is especially intriguing that in both Experiment 1 and Experiment 2, almost one-third of children chose to accept all trials, essentially providing no evidence of complex strategy use, while the remaining children evidenced at least some strategy use, but were distributed on a continuum of proficiency. The significance of the between-subjects analysis comparing the accuracy of 'accept' trials in Experiment 2 to the forced-choice recognition test in Experiment 3 suggests that some children in Experiment 2 who accepted all trials (those whose accuracy was very high) would not have benefited from using the 'decline' option, and thus may have been using a selective strategy. However, the performance of those children who accepted all trials but whose accuracy was low stands out from those children whose memory accuracy was low, but who were able to effectively utilize the 'decline' option to skip more difficult trials. Further exploration into the nature of individual differences that result in such disparate performance would be useful in elucidating the nature of complex cognitive functions in both normal and atypical populations of children.

This research is an important step in providing evidence that children can guide their behavior in complex ways using self-generated cues about their own knowledge structures, at least in a post-learning, retrieval task. To complete the picture, research needs to be done exploring whether infants can strategically direct their ongoing learning through monitoring knowledge states. Current research in our lab is under way to examine whether infant preference for novel or familiar stimuli predicts what they in fact know about those stimuli.

## References

- Aslin, R., Saffran, J., & Newport, E. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9* (4), 321–324.
- Butterfield, E., Nelson, T., & Peck, V. (1988). Developmental aspects of the feeling of knowing. *Developmental Psychology*, *24* (5), 654–663.
- Campbell, A., & Namy, L. (2003). The role of social-referential context in verbal and nonverbal symbol learning. *Child Development*, *74* (2), 549–563.
- Cultice, J., Somerville, S., & Wellman, H. (1983). Preschoolers' memory monitoring: feeling-of-knowing judgments. *Child Development*, *54* (6), 1480–1486.
- Diana, R., & Reder, L. (2004). Visual versus verbal metacognition: are they really different? In D. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 187–201). Cambridge, MA: MIT Press.
- Dunlosky, J., & Thiede, K. (1998). What makes people study more? An evaluation of factors that affect self-paced study. *Acta Psychologica*, *98*, 37–56.
- Fernandez-Duque, Baird, J., & Posner, M. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition*, *9*, 288–307.
- Flavell, J., Green, F., & Flavell, E. (2000). Development of children's awareness of their own thoughts. *Journal of Cognition and Development*, *1* (1), 97–112.
- Gerken, L. (2006). Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition*, *98*, B67–B74.
- Gomez, R. (2006). Dynamically guided learning. In Y. Munakata & M. Johnson (Eds.), *Attention and performance XXI. Processes of change in brain and cognitive development* (Vol. XXI, pp. 87–110). Oxford: Oxford University Press.
- Hampton, R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences*, *98* (9), 5359–5362.
- Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*, 341–348.
- Hunter, M., & Ames, E. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5*, 69–95.
- Hunter, M., Ames, E., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, *19* (3), 338–352.
- Inman, A., & Shettleworth, S. (1999). Detecting metamemory in nonverbal subjects: a test with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *25* (3), 389–395.
- Koenig, M., & Echols, C. (2003). Infants' understanding of false labeling events: the referential roles of words and the speakers who use them. *Cognition*, *87*, 179–208.
- Koenig, M., & Harris, P. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76* (6), 1261–1277.
- Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100* (4), 609–639.
- Koriat, A. (1994). Memory's knowledge of its own knowledge: the accessibility account of the feeling of knowing. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 115–135). Cambridge, MA: MIT Press.
- Kornell, N., Son, L., & Terrace, H. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science*, *18* (1), 64–71.
- Lockl, K., & Schneider, W. (2002). Developmental trends in children's feeling-of-knowing judgments. *International Journal of Behavioral Development*, *26* (4), 327–333.
- Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, *52*, 463–477.

- Nelson, T., & Narens, L. (1990). Metamemory: a theoretical framework and some new findings. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). New York: Academic Press.
- Nelson, T., & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.
- Newport, E., & Aslin, R. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, **48**, 127–162.
- Prins, F., Veenman, M., & Elshout, J. (2006). The impact of intellectual ability and metacognition on learning: new support for the threshold of problematicity theory. *Learning and Instruction*, **16**, 374–387.
- Reder, L., & Schunn, C. (1996). Metacognition does not imply awareness; strategy choice is governed by implicit learning and memory. In L. Reder (Ed.), *Implicit memory and metacognition* (pp. 45–77). Mahwah, NJ: Erlbaum.
- Schneider, W. (1998). Performance prediction in young children: effects of skill, metacognition and wishful thinking. *Developmental Science*, **1** (2), 291–297.
- Schneider, W. (1999). The development of metamemory in children. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: cognitive regulation of performance: interaction of theory and application* (pp. 487–514). Cambridge, MA: MIT Press.
- Schneider, W., Kron, V., Hunnerkopf, M., & Krajewski, K. (2004). The development of young children's memory studies: first findings from the Wurzburg longitudinal study. *Journal of Experimental Child Psychology*, **88**, 193–209.
- Schneider, W., & Pressley, M. (1997). *Memory development between 2 and 20* (2nd edn.). Mahwah, NJ: Erlbaum.
- Shields, W. (1999). *Nonverbal judgments of knowing by humans and rhesus monkeys*. Buffalo, NY: State University of New York at Buffalo.
- Shields, W.E., Smith, J.D., & Guttmanova, K. (2005). Confidence judgments by humans and rhesus monkeys. *Journal of General Psychology*, **132** (2), 165–186.
- Shimamura, A. (2000). Toward a cognitive neuroscience of metacognition. *Consciousness and Cognition*, **9**, 313–323.
- Smith, J.D., Minda, J., & Washburn, D. (2004). Category learning in rhesus monkeys: a study of the Shepard, Hovland, and Jenkins (1961) tasks. *Journal of Experimental Psychology: General*, **133** (3), 398–414.
- Smith, J.D., Shields, W.E., & Washburn, D. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences*, **6**, 317–373.
- Sodian, B., & Schneider, W. (1986). Metamemory–memory behavior relationships in young children: evidence from a memory-for-location task. *Journal of Experimental Child Psychology*, **45** (2), 209–233.
- Son, L., & Kornell, N. (2005). Metaconfidence judgments in rhesus macaques: explicit versus implicit mechanisms. In H. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 296–320). New York: Oxford University Press.
- Son, L., & Sethi, R. (2006). Metacognitive control and optimal learning. *Cognitive Science*, **30**, 759–774.
- Stipek, D., & Mac Iver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development*, **60**, 521–538.
- Thiede, K., & Dunlosky, J. (1999). Toward a general model of self-regulated study: an analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology*, **25** (4), 1024–1037.
- Thiessen, E., Hill, E., & Saffran, J. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, **7** (1), 53–71.
- Thiessen, E., & Saffran, J. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7- and 9-month-old infants. *Developmental Psychology*, **39** (4), 706–716.
- Washburn, D., Smith, J.D., & Tagliabate, L. (2005). Individual differences in metacognitive responsiveness: cognitive and personality correlates. *Journal of General Psychology*, **132** (4), 446–461.
- Wellman, H. (1977). Tip of the tongue and feeling of knowing experiences: a developmental study of memory monitoring. *Child Development*, **48** (1), 13–21.
- Zajonc, R. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, **9** (2), 1–27.

Received: 16 March 2007

Accepted: 10 September 2007