

Running Head: Metacognitive Monitoring and Control

Carving Metacognition at its Joints:

Protracted Development of Component Processes

Allison O’Leary and Vladimir M. Sloutsky

Department of Psychology, The Ohio State University

1835 Neil Avenue, Columbus, Ohio 43210

Word Count: 9753

Acknowledgments:

This research is supported by the NIH grant R01HD078545 and IES grant R305A140214 to VMS.

We thank Keith Apfelbaum and Heidi Kloos for helpful comments. Address correspondence to

Vladimir M. Sloutsky (sloutsky.1@osu.edu) or Allison O’Leary (oleary.83@osu.edu).

ABSTRACT

Two experiments investigated the development of metacognitive monitoring and control and conditions under which children engage these processes. In Experiment 1, 5-year-olds ($N = 30$) and 7-year-olds ($N = 30$), unlike adults ($N = 30$), showed little evidence of either monitoring or control. In Experiment 2, 5-year-olds ($N = 90$) were given performance feedback (aimed at improving monitoring), instruction to follow a particular strategy (aimed at improving control), or both. Across conditions, feedback improved children's monitoring, and instruction improved both monitoring and control. Thus, children's poor metacognitive performance likely reflects a difficulty engaging the component processes spontaneously, rather than a lack of metacognitive ability. These findings also suggest that the component processes are distinct, with both undergoing protracted development.

Carving Metacognition at its Joints: Protracted Development of Component Processes

Imagine a psychology instructor planning a broad survey class she has never taught before. To efficiently allocate her preparation time, she will need to assess her own knowledge. How much does she already know and how well? How difficult will it be to learn what is not yet known? And how long will it take? Once these questions are answered, she may decide to prepare her lectures in a way that maximizes quality (e.g., providing both deep and broad coverage) but minimizes time and effort (e.g., by focusing mainly on topics she is less familiar with). In short, she will need to access her own cognitive processes (such as memory and speed of learning) and use this information to guide or control her future behavior. These processes have been referred to as *metacognition* (Flavell, 1979; Nelson & Narens, 1990; Metcalfe & Shimamura, 1994).

Metacognition has been a subject of study for decades, with two complementary approaches emerging since Flavell (1979) coined the term. Some researchers have focused on metacognition as an *independent* variable that affects educational outcomes (see Hacker, Dunlosky, & Graesser, 2009). Within this approach, a primary question of interest has been how the development of metacognition improves the academic skills involved in reading, writing, math, and science.

The second approach (which is taken within the research presented here) has focused on metacognition as a *dependent* variable, with a primary focus on how people access their own cognition and how these abilities develop. These issues have been studied extensively in the context of efficient allocation of study time. In the classic experimental paradigm (see Son & Metcalfe, 2000, for a review), participants are asked to study two lists of word pairs for an upcoming memory test. Pairs on one list are related semantically (i.e., these are presumably easy to remember as pairs), whereas pairs on the second list are unrelated (i.e., these are presumably difficult to remember as pairs). Metacognitive ability is inferred from the different amounts of time participants spend studying the unrelated versus related word pairs. This paradigm suggests that metacognition hinges on two distinct sub-processes – monitoring and control. Specifically, noticing that one list is more

difficult to remember requires *monitoring*, whereas deciding to study the more difficult list for a longer time requires adjusting behavior accordingly (i.e., *control*). Whether the components are in fact distinct is an open question. One way to address this question is to examine the developmental time course of the components. Developmental asymmetries in the component processes (if found) would support the idea that the components are distinct.

There are a number of important findings pertaining to the development of metacognition and its components. These findings, however, present a somewhat conflicting developmental picture: whereas some findings suggest an early onset of metacognition, others suggest a late onset. Specifically, there are studies demonstrating evidence of monitoring and control in children as early as 3 years of age (Hembacher & Ghetti, 2014; Coughlin, Hembacher, Lyons, & Ghetti, 2014; Lyons & Ghetti, 2013), but there are also studies suggesting the onset of metacognition much later in life (Dufresne & Kobasigawa, 1989; Lockl & Schneider, 2004), with even adults experiencing difficulty in accessing their cognition (see Karpicke, Butler, & Roediger, 2009). In what follows, we review some of these findings.

Components of Metacognition and Their Development

Monitoring. Researchers have used multiple methods to study the ability to monitor one's cognition, including: ratings of confidence/uncertainty (see Lyons & Ghetti, 2011; Vo, Li, Kornell, Pouget, & Cantlon, 2014), judgments of performance (see Schneider, 1998), judgments of learning (JOL), feeling-of-knowing (FOK) judgments, judgments of difficulty, and ease-of-learning (EOL) judgments. These methods roughly fall into two categories. Some of the methods focus on *performance monitoring* (i.e., judging one's own performance on a task), whereas others focus on *task monitoring* (i.e., judging other aspects of the task, like difficulty or amount of effort required, without necessarily considering performance).

Judgments of performance, judgments of learning, feeling-of-knowing judgments, and confidence ratings fall into the category of performance monitoring -- an appraisal of one's success

in a task. To measure confidence, for example, participants are asked to report their certainty about a task response. There is evidence that children (and sometimes adults) tend to be over confident about their task performance (Roebbers, 2002), indicating that metacognition may be imperfect even in adulthood. However, even 3-year-olds report lower confidence for incorrect, relative to correct, responses (Lyons & Ghetti, 2011; Lyons & Ghetti, 2013; Hembacher & Ghetti, 2014). Hence, it has been concluded that even very young children can monitor their performance, at least under some circumstances.

Another category of metacognitive monitoring is task monitoring – judgments about the task itself or one’s potential (rather than actual) performance on the task. For example, judgments of task difficulty or of ease-of-learning may fall into this category. Task monitoring differs from performance monitoring in that it does not require an appraisal of actual performance, but rather an appraisal of some other aspect of the task (e.g., how difficult the task is, how much effort would be required to complete the task). Although people *may* rely on their past performance to assess these aspects of the task, they do not have to. They may instead assess the amount of (either actual or anticipated) effort required to perform the task, independent of performance. For example, adults avoid effortful tasks in an attempt to maximize performance *and* minimize effort (Kool & Botvinick, 2014; Kool, McGuire, Rosen, & Botvinick, 2010).

Similarly, in the study time allocation task, participants should notice that one type of trial (i.e., learning the unrelated word pairs) is more difficult than another. Importantly, they must do so without feedback regarding their performance. In this paradigm, children fail to monitor the difference in difficulty until about 6 years of age (Dufresne & Kobasigawa, 1989; Lockl & Schneider, 2004) – much later than the performance monitoring found in the studies mentioned above (Lyons & Ghetti, 2011; Lyons & Ghetti, 2013; Hembacher & Ghetti, 2014).

There are at least two possible explanations for the differential success in performance and task monitoring. First, it is possible that performance and task monitoring describe independent

aspects of metacognitive monitoring, which show asynchronous developmental trajectories (e.g., children demonstrate successful performance monitoring before successful task monitoring). Second, it is possible the tasks used by researchers to tap these components are responsible for these performance differences. For example, in studies showing early performance monitoring, children are probed to report on their certainty on *every trial*, which may prompt them to monitor their performance. This is in contrast to the study time allocation task, in which monitoring is only measured at the end of the task. It is possible that this repetitive probing improves children's performance monitoring through the course of the task. To avoid effects of continuous probing, the current study measured participants' monitoring in a batched fashion, only at the end of the task.

Control. Whereas metacognitive monitoring is the ability to represent information about the task at hand (including one's own performance), metacognitive control is the ability to use this information to adaptively adjust behavior to suit the task's demands. For example, to efficiently allocate study time, participants must use their knowledge (e.g., that one set is more difficult to remember than another) to formulate a strategy (e.g., that studying the difficult-to-remember pairs for a longer time is adaptive). Further, they must engage additional control processes to execute that strategy (i.e., by actually studying the difficult pairs longer, rather than showing no preference).

In the study time allocation task, despite being able to monitor the difference in difficulty by age 6, children younger than 8 years do not consistently study the difficult-to-learn items more (Dufresne & Kobasigawa, 1989; Lockl & Schneider, 2004). This suggests that monitoring and control are separable components and that proficient monitoring may develop before proficient control. However, more recent work has shown that children may exhibit evidence of metacognitive control even at 3-years-old, and that control is intimately tied to monitoring even at this age (Lyons & Ghetti, 2011; Lyons & Ghetti, 2013; Hembacher & Ghetti, 2014). These studies suggest that both monitoring and control develop early, and show similar developmental trajectories.

In an attempt to understand these divergent findings, we consider two differences in the

tasks used in these studies. First, in studies showing early metacognitive proficiency, children are *instructed* to withhold a response if they thought they had made a mistake (Hembacher & Ghetti, 2014). This provides an explicit strategy that children are encouraged to use throughout the task, obviating the need for children to *formulate* a strategy themselves. When the strategy is provided, children need only to *execute* it. In contrast, in the study-time allocation task, successful control depends on the child's ability to both formulate *and* execute a strategy. Many researchers have addressed children's difficulty with both (1) forming and selecting between strategies (see Reder, 1987; Siegler & Shipley, 1995; Siegler & Jenkins, 2014) and (2) behaviorally executing a chosen strategy (often referred to as production deficiency; see Kendler, 1972; Moely, Olson, Halwes, & Flavell, 1969).

Second, it is possible that, in studies reporting early metacognitive control, the frequency of monitoring probes matters. In other words, children may be more likely to withhold their responses after they have just expressed their uncertainty about each response verbally. This control behavior may be different from what children would do *spontaneously* (i.e., without frequent probing). As stated earlier, in the research presented here, children are only encouraged to explicitly reflect on their performance at the end of the task.

The discrepancy between findings suggesting early and late onset indicate that metacognition is not fixed and its deployment may be affected by how the task is structured. In the next section, we consider more deeply the influence of scaffolding (such as explicit strategy instruction) on early metacognition.

Effects of Scaffolding on Children's Metacognition

Both monitoring and control develop through childhood but, as reviewed above, the age at which children show metacognitive proficiency may depend on features of the task itself. For example, elementary-school children were more likely to use an organizational strategy to remember items when given explicit instruction about the utility of that strategy (Liberty &

Ornstein, 1973; Bjorklund, Ornstein, & Haig, 1977). This finding further suggests that children perform differently when provided with a strategy versus when having to formulate a strategy themselves, supporting the idea that task differences may be responsible for the discrepant findings described above (also see Destan, Hembacher, Ghetti, & Roebers, 2014 for a discussion of this possibility). If this is the case, we can predict that providing children with instruction to use a particular strategy will improve their metacognitive control by reducing the need to formulate a strategy spontaneously.

Fewer studies have focused on the role of scaffolding for metacognitive monitoring, but some suggest that receiving feedback about one's performance can lead to more accurate performance estimation (see Butler & Winne, 1995, for a review). Unlike adults, whose performance estimates often correlate with actual performance even in the absence of feedback (Yeung & Summerfield, 2012), children tend to overestimate their performance (Butler, 1990; Roebers, 2002). In the absence of an external cue regarding their performance, they must estimate or "self-generate" feedback to successfully monitor (Butler & Winne, 1995). Whether these kinds of estimations are accurate in childhood, however, is unclear, and the exact influence of trial-by-trial performance feedback on children's metacognition has not been tested directly. It is possible that explicit feedback will improve children's ability to monitor their behavior.

The Current Study

The current study had two primary aims. The first aim was to examine the development of both metacognitive monitoring and control. The second aim was to determine whether and how task characteristics, such as the presence of feedback or explicit strategy instruction, affect children's metacognitive performance. Achieving this aim would contribute to our understanding of conditions under which young children demonstrate proficient metacognition.

In Experiment 1, we examined how 5- and 7-year-olds and adults engage each component of metacognition spontaneously, when given neither feedback about performance nor instructions as to

how to perform optimally. We chose this age range because it (a) covers most of the ages of the putative onset of metacognitive proficiency reported in previous studies and (b) even the oldest children are still developing top-down control processes that are likely linked to metacognitive control (Davidson, Amso, Anderson, & Diamond, 2006). We also included an adult sample to identify components of metacognition that change between childhood and adulthood.

To address the second aim, we conducted Experiment 2 to investigate the effects of feedback and instruction scaffolding on children's metacognitive monitoring and control. In addition to examining *whether* instruction scaffolding can have systematic effects on metacognition, we also tested *how* these effects transpire. For example, if feedback improves children's metacognitive monitoring, but not their control, this would provide evidence for independence of monitoring and control processes. However, if improvements in monitoring result in improvements in control, and vice versa, this would provide evidence for interdependence. We further discuss these predictions in the introduction of Experiment 2.

We predicted that performance feedback would improve children's metacognitive *monitoring* by providing an external cue to their performance. We also predicted that strategy instruction would improve children's ability to successfully *control* their behavior by eliminating the need to formulate a strategy spontaneously. Finally, we predicted improvements in both monitoring and control when children are provided with both feedback *and* strategy instruction.

To adequately address these aims and examine the development of metacognition, it was important to use a base-level task (i.e., the task that would generate *meta*-level behavior) in which performance would be comparable amongst children and adults. We chose to use a simple numerical discrimination task. In this task, participants saw two sets of dots, and were asked to judge which of the sets had contained more dots. Participants were exposed to trials at two levels of difficulty (1:2 ratio vs. 9:10 ratio, on average), and each ratio perfectly corresponded to a particular color in which the dots would be displayed.

At the beginning of each trial, participants were allowed to choose the ‘game’ they would play on that trial by selecting the color of the dots in the game (i.e., red or blue). Because participants were incentivized to perform as accurately as possible, we used the proportion of “easy” trials chosen as our measure of *metacognitive control*. Crucially, participants were not told which game (red or blue) was easier, and had to discover this through experience with the games. Therefore, discovering which game was more difficult should require either performance or task monitoring. Furthermore, adjusting their choices to maximize performance and minimize effort would require the deployment of control.

This measure of control is based on a long tradition of research demonstrating that humans are “cognitive misers” in their tendency to minimize mental effort (see Kool, et al., 2010, for an extensive review). Thus, when given the opportunity to choose between an easy and a difficult task, the most adaptive strategy would be to choose the easier option and avoid the more effortful one that also results in lower performance. In our task, participants had to generate a strategy (e.g., “if I choose the blue game, I make fewer mistakes”) and execute it (by choosing the easier game) to perform optimally. To our knowledge, this is the first study to systematically address whether children, like adults, are “cognitive misers.”

We also measured participants’ performance and task monitoring at the end of the experiment. Participants were asked to estimate the proportion of their correct responses throughout the task (*performance monitoring*), which we could then compare to their actual performance to assess the accuracy of their estimations. Further, we asked participants to indicate (1) whether they noticed the tasks’ differential difficulty and (2) which task was easier (*task monitoring*).

EXPERIMENT 1

In Experiment 1, we focused on developmental differences in the monitoring and control components of metacognition in a task that required participants to engage those components spontaneously.

Method

Participants

A sample of 5-year-olds ($N = 30$, 15 girls, $M = 5.43$ years, $SD = 0.25$ years), 7-year-olds ($N = 30$, 15 girls, $M = 7.51$ years, $SD = 0.27$ years), and undergraduate students from The Ohio State University ($N = 30$, 14 women, $M = 21.97$ years, $SD = 5.02$ years) participated in this experiment. Children were recruited through local daycares, preschools, and elementary schools located in Columbus, Ohio. Undergraduate students received course credit for their participation.

Materials and Design

Stimuli were presented using OpenSesame presentation software (Mathôt, Schreij, & Theeuwes, 2012) on either a Dell PC (for adults) or a Dell laptop accompanied by a touch screen (for children). Stimuli in the numerical discrimination task consisted of sets of dots presented in pairs. There were two levels of discrimination difficulty: *easy* and *difficult*. Easy discriminations included a 1:2 ratio of dots and were instantiated with the following sets: 4 vs. 8, 5 vs. 10, 6 vs. 12, 7 vs. 14, 8 vs. 16, 9 vs. 18, 10 vs. 20, 11 vs. 22, 12 vs. 24, and 13 vs. 26. The difficult discriminations included sets that had a 9:10 ratio or smaller and were instantiated with the following sets: 9 vs. 10, 10 vs. 11, 11 vs. 12, 12 vs. 13, and 13 vs. 14. Previous research has demonstrated that these two ratios are differentially difficult to discriminate for both children and adults (Halberda & Feigenson, 2008).

For each participant, each level of difficulty was randomly assigned to a separate color at the beginning of the experiment. Therefore, for some participants the dots were blue in easy discriminations and red in difficult discriminations, whereas for others the reverse assignment was used. Importantly, the color-difficulty contingency was stable within participants, but varied randomly across participants. Figure 1 shows the trial sequence. Each trial consisted of a choice opportunity, fixation, test stimulus, and response screen.

Procedure

Before the experiment began, all participants were incentivized to complete the task as accurately as possible. Participants were instructed that the object of the game was to correctly discriminate quantities of dots. Adults were told that they would earn 5 points for each correct answer, and that they would lose 5 points for each incorrect answer or if they did not respond in the time allotted. Their goal was to accumulate as many points as possible. Children were told that they would acquire a point for each correct answer, and would lose a point for each incorrect answer or if they did not respond to a trial in time. They were told that the more points they received, the more stickers they could select at the end of the task. Points were not actually tabulated throughout the experiment and all children received the same number of stickers.

Measuring Control. Prior to each discrimination trial, participants were allowed to choose that trial's difficulty by selecting between the two corresponding dot colors. Importantly, participants were not instructed that the color was related to the task difficulty, nor which task was easier, and had to learn the color to level-of-difficulty contingency through experience with the task. During each choice opportunity, participants were presented with a red and a blue dot, whose placement on the left or right side of the screen were randomized on each trial. They were allowed to choose to play either the "red game" or the "blue game" by clicking or touching the appropriate dot. Assuming that people tend to maximize reward and minimize effort (Kool et al., 2010), the proportion of easy task choices should reflect the tendency to control behavior.

Measuring Discrimination Performance. Following the participant's choice, a white circle fixation target appeared in the center of the screen for 500 milliseconds. Then the test stimulus appeared, which consisted of two grey boxes each containing a randomly positioned array of dots in the color the participant had just chosen. The number of dots in each array was presented according to the ratio associated with the chosen color (one color corresponded to easy to discriminate ratios, whereas another to difficult). These dot arrays were shown for 500 milliseconds. Finally, the dots disappeared leaving only the empty grey boxes, and participants were asked to indicate which of the

two boxes had contained more dots. The boxes remained on screen until the participant made a response, or until 7000 milliseconds had passed. Adults indicated their response using a computer mouse, whereas children made their response by touching the selected box on a touchscreen. All participants completed two practice trials, followed by 30 test trials. Importantly, the proportions of easy and difficult discrimination trials for each participant depended on their choices during each choice opportunity.

Measuring Monitoring. Following the test trials, we assessed participants' performance and task monitoring. To evaluate *performance monitoring*, we asked participants to estimate (on a scale of 1 to 5) the proportion of trials they had answered correctly. Children were asked how many trials they had gotten correct from the following options: none of them, some of them, half of them, most of them, or all of them. They indicated their answer by selecting a corresponding circle that was 0%, 25%, 50%, 75%, or 100% filled. Adults were asked to select the proportion (from 0%, 25%, 50%, 75%, and 100%) that best corresponded to the proportion of trials answered correctly. This allowed us to measure participants' "absolute" performance monitoring, or how accurately they estimated their overall performance. After this, participants were asked how many trials of each color they had answered correctly (e.g., "How many of the [red/blue] ones did you get correct?") in the same manner. The order of these two questions was randomized. This provided a measure of participants' "relative" performance monitoring, in that we could assess whether they rated their performance higher for easy than for difficult trials.

At the end of the task, three questions were used to assess participants' *task monitoring*. First, we asked participants whether they thought the red game and the blue game were the same or different. If they answered 'same,' the experiment terminated. If they answered 'different,' they were asked whether they thought one game was easier than the other. If they answered 'no' to this question, the experiment terminated. If they answered 'yes,' they were asked which game they thought was easier. At the end of the experiment, all adults were told that their performance was

“excellent,” and all children were awarded 3 stickers (as is customary in our lab, and did not reflect an additional reward for performance).

Results and Discussion

Preliminary Analyses. For 5- and 7-year-olds, there was no effect of sex on any of our measures (all $ps > .08$). There was no effect of sex on adults’ monitoring performance, whereas males outperformed females on our measure of metacognitive control ($p = .032$). This finding, however, is difficult to interpret and does not inform our questions of interest, so we collapsed across sex for all the following analyses.

Discrimination Accuracy and Response Times. Before proceeding with the main analyses, it was necessary to validate that the two discrimination tasks were in fact differentially difficult for both children and adults. Indeed, as shown in Table 1, participants of all age groups were significantly more accurate in the “easy” discrimination task than the “difficult” task, all $ts > 8.8$, $ps < .001$, $ds > 3.3$. The average difference in accuracy in the easy and difficult tasks was 29% ($SD = .17$) for adults, 31% ($SD = .13$) for 7-year-olds, and 31% ($SD = .15$) for 5-year-olds, which were not significantly different, $F(2, 86) = .158$, $p = .854$, $\eta^2 = .003$. This finding is important because it means the difference in difficulty was comparable across age groups. Thus, any reported differences in metacognition cannot be due to differences in base-level task performance.

In addition, as shown in Table 1, adults’, 7-year-olds’, and 5-year-olds’ response times were significantly slowed in the difficult task relative to the easy task, all $ts > 3.25$, $ps < .005$, $ds > 1.2$. These results are worth noting – they suggest that even young children implicitly detected the difference in difficulty, slowing their responses to difficult trials. *Metacognitive Control.* To assess metacognitive control, we examined how often participants chose the less demanding, easy task (see Figure 2). As predicted, adults chose the easy task more frequently than would be expected by chance, $M = 75\%$, $t(29) = 7.75$, $p < .001$, $d = 2.88$. In addition, as shown in Figure 3, adults’ choices of the easy task increased with task experience, as evidenced by the effect of block (each containing

6 trials), $F(4, 116) = 4.85, p < .005, \eta^2 = .143$. This increase exhibited a linear trend, $F(1, 29) = 8.45, p < .01, \eta^2 = .226$. Neither 5-year-olds ($M = 50.8\%$) nor 7-year-olds ($M = 49.2\%$) chose the easy task consistently ($ts < 1, ps > .6, ds < .2$), with both age groups choosing the easy task less often than adults, $F(2, 87) = 30.09, p < .001, \eta^2 = .41$ (see Figure 2).

We also examined the proportion of individuals who systematically chose the easy task. If a participant chose the easy task on at least 20 out of 30 trials ($p < .05$, according to binomial probability), they were considered an “optimizer.” Twenty-one adults (70% of the sample) optimized by systematically choosing the easy task. A single 5-year-old (3% of the sample, and a single 7-year-old (3% of the sample) were classified as optimizers. All other children simply switched between the two games. The proportions of child optimizers were significantly smaller than the proportion of adult optimizers, $X^2(2, N = 90) = 46.72, p < .001$. Taken together, these findings indicate that only adults exhibited evidence of optimizing their performance and minimizing effort.

Task Monitoring. We asked participants three questions: (1) whether the two games were different, (2) whether one game was easier, and (3) which game was easier. To evaluate participants’ task monitoring, we calculated a composite score with a maximum of two points. If they correctly indicated that one game was easier than the other, they received a point. If they then correctly identified which of the two games was easier, they received a second point. Participants who failed to notice any difference between the two tasks did not receive a score. Adults’ average composite task monitoring score was 1.88 out of a possible 2, indicating that they consistently tracked the difference in difficulty. In contrast, 5- and 7-year-olds’ scores of .90 and 1.09, respectively indicated that children struggled to monitor task difficulty. Whereas a one-way ANOVA revealed a significant difference between the performance of children and adults, $F(2, 66) = 11.56, p < .001, \eta^2 = .26$, the two groups of children were not significantly different from one another, $p = .42$.

We also identified the proportion of individuals who answered all 3 questions correctly (i.e., those who showed highly proficient monitoring). Twenty-three out of 30 adults (77% of the sample) correctly identified which game was easier (i.e., answered all three questions correctly). Interestingly, the majority of these adults (i.e., 17 out of 23) were consistent optimizers. Overall, more adults proficiently monitored the task than children, $\chi^2(1, N = 90) = 22.81, p < .001$. Eleven of 30 7-year-olds (37% of the sample) correctly answered all three questions, but only 1 of those 11 chose the easy task systematically. Only five of 30 5-year-olds (17% of the sample) correctly answered all three questions, and only one of these five was an optimizer. Overall, more 7-year-olds than 5-year-olds answered all three questions correctly, although this difference was marginally significant, $p = .08$.

Given their low monitoring scores, it is possible that children failed to exhibit control and choose the easier task simply because they failed to learn the contingency between the color and task difficulty. To determine whether children who successfully monitored were more likely to control their behavior, we compared the proportion of easy task choices of children who successfully and unsuccessfully monitored the task (i.e., noticed which game was easier). There were no differences in the control performance of these two groups, $p > .97, d = .01$. Therefore, even those children who successfully learned the contingency did not reliably select the easier game – successful task monitoring did not necessarily lead to successful control.

Performance Monitoring. We asked participants to estimate the proportion of correct responses on all trials, on only red trials, and on only blue trials. We used these questions to assess (1) their sensitivity to absolute accuracy (i.e., how precise their estimation was), (2) the direction of their estimations (i.e., whether the sample over- or underestimated performance), and (3) their sensitivity to relative accuracy (i.e., whether they noticed that they performed more accurately on easy relative to difficult trials).

To assess participants' sensitivity to their absolute accuracy, we first calculated the absolute

value of the difference between their estimated and actual accuracy. However, this value is biased in favor of individuals whose accuracy happened to be in the middle of their chosen interval. For example, if participant A chose the interval corresponding to 50%, and actually completed 50% of trials correctly, their value would be 0. If participant B also chose the interval corresponding to 50%, and actually completed 40% of trials correctly, their value would be 10 despite the fact that they chose the most appropriate interval. To avoid this bias, we adjusted these values to suit our use of a discrete scale. Because our measure used intervals of 25%, if participants' estimations were within 12.5% of their actual accuracy (i.e., within that interval), we transformed their difference score to 0. If participants' estimations differed by more than 12.5% from their actual accuracy, we subtracted 12.5% from their actual difference score. Thus if a participant completed 40% of trials correctly and chose the interval corresponding to 50%, their difference score was 0. However, if a participant completed 35% of trials correctly and chose the interval corresponding to 50%, their difference score was equal to 2.5 (i.e., 15-12.5). Difference scores of 0 indicated accurate estimates, whereas scores greater than 0 indicated misestimated performance.

Adults' average absolute performance monitoring score was 12.8%, which was significantly different from 0, $t(29) = 5.88$, $p < .001$, $d = 2.18$, indicating that adults' estimations were imprecise. Figure 5 displays the *direction* of participants' performance estimations (i.e., *unadjusted* estimated – actual accuracy). Positive numbers indicate that participants overestimated their performance, whereas negative numbers indicate that they underestimated performance (values around 0 indicate that participants had accurate estimates). Most of adults' scores were below zero, indicating that they systematically underestimated their performance.

Five-year-olds' average absolute performance monitoring score was 11.2%, indicating that they also misestimated performance, $t(29) = 5.52$, $p < .001$, $d = 2.05$. Although they were not different from adults in terms of this overall effect, the *direction* of the effect did differ: whereas adults tended to underestimate their performance, 5-year-olds tended to overestimate (see Figure 5).

Seven-year-olds, with a difference score of only 5%, estimated their performance more precisely than both 5-year-olds and adults, $F(2, 87) = 4.56, p < .05, \eta^2 = .10$, although this score was still different from zero, $t(29) = 3.39, p < .005, d = 1.26$. These results suggest that adults focused more on errors (thus underestimating their performance) and 5-year-olds focused more on their successes (thus overestimating their performance). This interpretation explains the pattern of results and suggests that 7-year-olds are a transitional group (perhaps these participants focused on both errors and correct responses).

Finally, to evaluate participants' sensitivity to relative performance, we calculated the percentage of individuals who correctly rated their accuracy in the easy task higher than that in the difficult task. Twenty-five adults (83% of the sample), 15 7-year-olds (50% of the sample), and 11 five-year-olds (37% of the sample) correctly rated their accuracy on easy trials as higher. More adults correctly noticed the difference in their performance than 5-year-olds and 7-year-olds, $\chi^2(1, N = 90) = 14.12, p < .005$. Nineteen of these 25 adults were optimizers. None of these 15 7-year-olds, and only 1 of these 11 five-year-olds, were optimizers. *Summary of Findings*

Experiment 1 demonstrated developmental differences in the metacognitive monitoring and control of 5-year-olds, 7-year-olds, and adults. Adults (1) accurately monitored the difference in difficulty between the two tasks and (2) minimized effort and optimized performance by choosing the easier of the two tasks. Furthermore, most adults correctly rated their accuracy on easy trials as higher than that on difficult trials.

Five-year-olds, on the other hand, showed immaturities in both monitoring and control. Children failed to consistently use a strategy to control their behavior (i.e., they chose the easy and difficult tasks equally often). Further, less than one fifth of the 5-year-olds answered all three task monitoring questions correctly, and only about a third of 5-year-olds reported having higher accuracy in the easy game. Hence, in contrast to adults, the majority of these children failed to monitor either their own performance or the differential task difficulty. Even those children who did

successfully monitor (i.e., those who could identify which game had been easier) did not attempt to optimize their performance. This suggests that a trivial explanation of the findings (i.e., that children's control failure occurred simply because they failed to learn the contingency between the color and task difficulty) was not the case.

In our measures of monitoring, 7-year-olds appear to be a transitional group. Although significant differences between 5- and 7-year-olds only transpired in terms of their performance monitoring, 7-year-olds had somewhat higher scores on our measures of task monitoring as well. However, neither 5-year-olds nor 7-year-olds adopted the optimal strategy of choosing the easier task, despite the fact that over a third of 7-year-olds correctly identified the easier task. This suggests that, unlike adults, children were not "cognitive misers" -- they did not spontaneously avoid a challenging task. These findings also suggest that monitoring can develop without subsequent increases in control, supporting the idea that monitoring and control are dissociable and show different developmental trajectories.

These trajectories point to some differences with previous work. For example, it has been found that children as young as 3-years-old are capable of monitoring their performance in a task. However, we have described poor monitoring ability in 5-year-olds in terms of both performance and task monitoring, which is more similar to the trajectory seen in studies of study-time allocation. We think this difference transpired, at least in part, due to the differences in the tasks described in the introduction. In Experiment 2, we investigate the possibility that specific task features can determine whether children engage in metacognitive processes.

EXPERIMENT 2

Experiment 1 required children to monitor and control their behavior *spontaneously* (i.e., with no performance feedback or instruction regarding an optimal strategy). However, there are reasons to believe that children's monitoring and/or control ability may transpire when external scaffolding is provided. For example, in studies showing early monitoring and control, children (1)

were cued to appraise their performance (i.e., asked to make an explicit confidence judgment) on every trial, and (2) were provided with a strategy for controlling behavior (e.g., to put the answer in the “closed eyes” box to avoid making a mistake). By asking children to appraise their performance on every trial, the researchers prompted children to reflect on their performance. This prompting may make it easier for children to monitor their performance, thus resulting in the observed early monitoring proficiency. We think that performance feedback also provides external cues about one’s performance and may have similar effect on children’s monitoring (Butler & Winne, 1995).

On the basis of these considerations, we hypothesized that performance feedback would improve children’s metacognitive monitoring. Further, if the monitoring and control components are dissociable, feedback may facilitate children’s performance monitoring, but not necessarily their control processes. We also hypothesized that instruction (or strategy scaffolding) would improve children’s control processes by eliminating the need to spontaneously formulate a strategy. If children in Experiment 1 failed to optimize performance due to immature control processes, children should not optimize even in the presence of instruction. Conversely, if children do benefit from instruction, this would indicate that what develops is their ability to spontaneously *formulate* a strategy. Further, facilitation of only control, but not monitoring, under the instruction condition, would provide evidence for the dissociability of the two components.

In Experiment 2, we investigated the effects of feedback only, strategy instruction only, and the compound effects of feedback and instruction on 5-year-olds’ metacognitive monitoring and control. We included only 5-year-olds in this experiment to be able to make more direct comparisons to studies reporting early metacognitive proficiency.

Method

Participants

Ninety five-year-olds participated in this experiment: 30 in the Feedback Only condition (12 girls, $M = 5.37$ years, $SD = .23$ years), 30 in the Instruction Only condition (11 girls, $M = 5.23$

years, $SD = .17$ years), and 30 in the Feedback + Instruction condition (13 girls, $M = 5.42$ years, $SD = .25$ years). Children were recruited through local daycares and preschools in Columbus, Ohio.

Materials, Design, and Procedure

Stimuli and procedure were similar to those used in Experiment 1, except that participants were provided with performance feedback, given instructions, or both. In the Feedback Only condition, participants received performance feedback after each discrimination response. They were told that if they correctly chose the box containing more dots, a smiley face would appear and they would hear a high tone. Participants were also told that if they responded incorrectly, or if they did not respond within 7000ms, they would see a sad face and hear a low tone.

In the Instruction Only condition, participants were told that one game was easier than the other, and were reminded before each trial to remember the task's "magic rule:" to choose the easier game. Importantly, they were not told *which* game was easier, and still had to discover this through experience with the task.

In the Feedback + Instruction condition, children received performance feedback after every trial. They were also told that one game was easier and were reminded before each trial to follow the task's "magic rule:" to choose the easier game.

Results and Discussion

Preliminary Analyses. There was no effect of sex on participants' performance in any of our measures (all $ps > .07$), so we collapsed across sex in all the following analyses.

Discrimination Accuracy and Response Times. As in Experiment 1, children were more accurate in the easy task than the difficult task, in all conditions (all $ts > 5.2$, all $ps < .001$). In addition, 5-year-olds responded more slowly to difficult trials than easy trials in both the Feedback Only and Feedback + Instruction conditions (both $ts > 2.32$, both $ps < .01$). Although children in the Instruction Only responded more slowly to difficult trials numerically, this difference did not reach significance, $p = .264$. See Table 2 for discrimination accuracy and response times for each

condition.

Metacognitive Control. To test the effects of feedback and instruction, we considered data from the 5-year-olds in both Experiments 1 and 2. This gave us a fully crossed design, with Experiment 1 serving as a no-Feedback and no-Instruction baseline and the three conditions of Experiment 2 introducing Feedback only, Instruction only, and both Feedback and Instruction. This design allowed us to conduct a 2 (no feedback vs. feedback) x 2 (no instruction vs. instruction) ANOVA on children's proportion of easy task choices (see Table 2). This analysis revealed a main effect of instruction on the proportion of easy task choices $F(1, 116) = 5.57, p < .05, \eta^2 = .05$, as predicted. Children's metacognitive control improved when provided with instruction to choose the easy task. There was no effect of feedback, $F(1, 116) = 1.28, p = .26, \eta^2 = .01$, and no significant interaction, $F(1, 116) = .618, p = .43, \eta^2 = .01$. Although the interaction was not significant, it is worth noting that 5-year-olds chose the easy task reliably above chance in the Feedback + Instruction condition only (61%, $t(29) = 2.87, p < .01, d = 1.07$). We also assessed the proportion of optimizers (i.e., individual children who chose the easier task on at least 20 trials) in each condition (see Table 2). More children optimized when provided with additional instruction (i.e., comparing the conditions where participants received instruction with those where they did not), $X^2(1, N = 120) = 11.76, p < .005$. However, feedback did not affect the proportion of optimizers (i.e., comparing the conditions where children received feedback with those where they did not), $X^2(1, N = 120) = .960, p = .327$. In the Baseline, 1 child was an optimizer (3% of the sample), in the Feedback Only condition, 2 children were optimizers (7% of the sample), in the Instruction Only condition 7 children were optimizers (23% of the sample), and in the Feedback + Instruction condition 10 children were optimizers (33% of the sample). Therefore, providing an explicit strategy improved individual children's metacognitive control, whereas performance feedback did not.

Task Monitoring. How did feedback and instruction impact task monitoring across the four conditions? As in Experiment 1, we calculated a composite score of children's task monitoring. A 2

(no feedback vs. feedback) x 2 (no instruction vs. instruction) ANOVA revealed significant main effects of both feedback, $F(1, 96) = 4.61, p < .05, \eta^2 = .05$, and instruction, $F(1, 96) = 8.64, p < .005, \eta^2 = .08$, on 5-year-olds' task monitoring scores, with no significant interaction, $p > .7$. Five-year-olds' task monitoring scores were .90 in the Baseline condition, 1.19 in the Feedback Only condition, 1.30 in the Instruction Only condition, and 1.69 in the Feedback + Instruction condition. Therefore, whereas only instruction improved metacognitive control, both instruction and feedback resulted in improved task monitoring.

Performance Monitoring. Children's absolute performance monitoring scores were calculated in the way described in Experiment 1. There were no effects of feedback or instruction on children's performance monitoring scores, all $ps > .33$. The average adjusted difference score was 11.2% in the Baseline condition, 11.7% in the Feedback Only condition, 13% in the Instruction Only condition, and 9.6% in the Feedback + Instruction condition. Five-year-olds tended to overestimate their performance, showing underestimation only in the Feedback Only condition. We also examined relative performance monitoring -- the proportion of children who rated their performance as higher on easy relative to difficult trials. The proportion of correct responders did not differ as a function of feedback or additional instruction, both $ps > .8$). Therefore, unlike task monitoring, children's performance monitoring was unaffected by either feedback or instruction.

Summary of Findings

Across the four conditions, 5-year-olds exhibited evidence of metacognitive control when they were provided with a strategy (i.e., to choose the easier task). These findings suggest that the mechanisms underlying metacognitive control are not completely immature at this age. Instead, poor performance in Experiment 1 likely stemmed from a failure to engage the processes spontaneously. Providing a strategy in some conditions of Experiment 2 facilitated the engagement of control processes by obviating the need to formulate a strategy (the only remaining demand was to execute the strategy appropriately).

In addition, as predicted, feedback affected children's task monitoring: receiving external feedback about performance helped children recognize which task was easier. Instruction also improved children's task monitoring, indicating that children were better at identifying which task was easier when prompted to choose the easy game. Providing children with a strategy likely encouraged them to monitor their progress toward that strategy in a way they would not have spontaneously. Surprisingly, in contrast to task monitoring, there was no effect of feedback on children's performance monitoring: regardless of whether or not feedback was provided, children tended to overestimate their overall performance. In addition, the majority of children did not provide accurate estimates of whether their performance was higher in the easy or in the difficult task. Although more research is needed to further examine the unresponsiveness of performance monitoring to feedback, these findings suggest that task monitoring and performance monitoring are potentially independent and may exhibit different developmental trajectories.

GENERAL DISCUSSION

This study addressed two main issues. First, we investigated the development of the monitoring and control components of metacognition. We found that adults spontaneously monitored the task and optimized behavior, whereas 5-year-olds and 7-year-olds failed to do so. Second, we examined effects of feedback and strategy instruction on children's metacognitive performance. It was found that providing 5-year-olds with a strategy increased their ability to control behavior and select an easier task. In addition, both feedback and strategy instruction improved 5-year-olds' task monitoring.

Spontaneous Monitoring and Control Across Development

Experiment 1 revealed a number of important developmental differences. Consistent with their status as "cognitive misers," adults chose an easier task to optimize performance and reduce effort. In contrast, neither 5- nor 7-year-olds controlled their behavior by using an adaptive strategy. This was not due to a lack of task monitoring ability – children who were well aware of which task

was more difficult were no more likely to control behavior than those who were not. Therefore, at the very minimum, in contrast to adults, young children are not cognitive misers. As discussed below, we believe that that this reflects their inability to formulate and implement a strategy that would minimize effort.

In addition, when asked to estimate their performance, 5-year-olds tended to overestimate, whereas adults tended to underestimate. This finding suggests that, in the absence of feedback (i.e., an external error signal), adults tend to focus more on possible errors, whereas children tend to focus on possible successes. Seven-year-olds may represent a transitional group for performance monitoring, having overcome their overestimation bias, but having not yet developed an underestimation bias. These developmental differences are consistent with work demonstrating that (a) adults show greater changes in neural activity subsequent to making an error than do young children (Wiersema, van der Meere, & Roeyers, 2007) and (b) better monitoring of errors in adults is associated with more adaptive adjustments to behavior (e.g., Holroyd & Coles, 2002). It is possible that adults optimized their performance more readily than children due to their higher sensitivity to errors.

These findings point to a protracted development of both metacognitive monitoring and control, somewhat consistent with studies of children's study time allocation. How can we reconcile the poor metacognitive monitoring and control in the 5- and 7-year-olds reported here with studies showing metacognitive proficiency as early as 3 years of age? We addressed this question in Experiment 2 by manipulating the presence of task variables that could contribute to precocious metacognitive performance.

Effects of Scaffolding on Children's Metacognition

We hypothesized that performance feedback would improve children's metacognitive monitoring by providing an external cue to their performance. We also hypothesized that providing children with a strategy would improve their metacognitive control performance by eliminating the

need to formulate a strategy oneself. In this case, children need only execute the strategy to optimize performance.

Monitoring. As predicted, 5-year-olds' task monitoring improved when children were given feedback about their performance. This indicates that even 5-year-olds have the ability to monitor task difficulty, although they may rely on an external cue like performance feedback to successfully do so. Children's task monitoring also improved when they were provided with an adaptive strategy (i.e., to choose the easier task). Perhaps cueing children to the fact that one game would be easier prompted them to identify which of the two games was easier. Thus, not only did children's task monitoring improve with an external cue (e.g., feedback), it also benefited from a more "top-down" cue (e.g., the goal to choose the easier game). Like the frequent monitoring probes used in studies showing early metacognition (Lyons & Ghetti, 2011; Lyons & Ghetti, 2013; Hembacher & Ghetti, 2014), these "top-down" cues may have improved monitoring of difficulty by directing one's attention to the task.

Children's estimations of their performance, however, were not affected by either feedback or instruction. This suggests that the ability to precisely estimate one's performance and the ability to judge a task's difficulty are separable. To judge a task's difficulty, as in our task monitoring measure, one can monitor a number of cues to judge the differential difficulty (e.g., whether they expended less effort, made fewer mistakes, or received more smiley faces on one task relative to the other). To accurately estimate performance, however, children must *precisely* keep track of the proportion of correct responses across time. Given that performance estimations did not improve with feedback, this suggests that improvements in children's task monitoring did not transpire through improvements in performance monitoring. One interpretation is that feedback improves children's task monitoring by altering their perceptions of expended effort. In other words, receiving negative feedback might translate to an appraisal like, "this game is hard," rather than "I am doing badly in this game." Future work should further unpack the distinction between task and

performance monitoring in children as well as adults, for whom we expect a similar pattern of results to transpire.

Control. Unlike their task monitoring, 5-year-olds' metacognitive control was unaffected by the presence of feedback. This suggests that improvements in monitoring need not lead to improvements in control, supporting the idea that monitoring and control are independent processes. Children's control did improve, however, when they were provided with a strategy (i.e., to choose the easier game). This suggests that children's control processes are also somewhat functional, but that children struggle to initiate them spontaneously. Instructing children to choose the easier game eliminated the need for children to formulate a strategy, which increased their successful strategy execution. These findings are consistent with classical work on children's strategy use, which suggest that children often do not formulate *and* implement adaptive strategies spontaneously until late childhood (Moely et al., 1969).

The difference between strategy formation and strategy execution can also be observed in other developmental tasks. For example, in the dimensional change card sort (DCCS) task, children are asked to sort cards by one dimension (e.g., color) and, after a number of trials, to switch and sort by a conflicting dimension (e.g., shape). Notably, young children often fail to execute an appropriate response (i.e., to sort according to the new dimension) despite being able to verbally report the correct strategy (Zelazo, Frye, and Rapus, 1996). Therefore, it is likely that strategy formation and execution are separable components. Future work should assess the development of these processes and their unique contributions to metacognitive control.

It is worth noting that, although children did successfully control behavior with scaffolding, these effects were not overwhelming. In all the conditions, however, 5-year-olds (as well as 7-year-olds in Experiment 1) slowed their response times to difficult relative to easy discrimination trials. Thus, children may have had a more implicit representation of the task difficulty that did not directly translate to successful explicit monitoring or more overt behavioral control. Previous work

has highlighted the possibility of “implicit metacognition,” or metacognition which occurs outside of conscious awareness (Reder, 1996). It is possible that this differential slowing (similar to post-error slowing) highlights implicit control processes. Under this explanation, children may not need to explicitly monitor which task was more difficult to control behavior accordingly. Future work should investigate the relation between implicit and explicit indicators of control, including whether one gives rise to the other.

What Develops and Why?

These findings make it clear that, throughout development, humans experience changes and improvements in performance monitoring, task monitoring, and metacognitive control (including both strategy formation and execution). So far, we have described these changes, but why do these changes occur?

Changes in performance monitoring are characterized by a tendency to overestimate performance in early childhood, followed by a tendency to underestimate in adulthood. If it is difficult to estimate performance accurately, underestimation is more likely to lead to strategy adjustments than overestimation. Perhaps children’s overconfidence obviates their perceived need to search for a more adaptive strategy, whereas decreasing overconfidence with age may stimulate the need to search for a more adaptive strategy. One possible source of developmental change is the ability to self-generate error signal when performance is uncertain. An additional source of development is the ability to accumulate error signal across trials and keep this record in working memory. The memory mechanism used to track successes and failures across trials in this task, as well as similar tasks (e.g., the Less is More task; Carlson, Davis, & Leach, 2005), remains unclear.

We also found developmental improvements in task monitoring ability, in that adults were able to indicate which task was easier even in the absence of feedback, whereas children needed feedback to do so. This suggests that part of what develops is the ability to self-generate feedback, which may be dependent on experience receiving feedback in the early school years. This idea is

supported by the fact that 7-year-olds monitored the task somewhat better than 5-year-olds. Perhaps receiving feedback (e.g., on assignments, on answers given in class, etc.) over time helps children draw conclusions about whether they excel at a task or not. Over time, children should be able to draw these conclusions on the basis of internal cues like increased effort, which is often associated with negative feedback as one is acquiring a new skill.

What drives the development of the control component, given its relative independence from monitoring? First, young children may have difficulty maintaining the current goal (in this case, to optimize performance) in working memory (Marcovitch, Boseovski, Knapp, & Kane, 2010). This ability undergoes development during the early childhood years, likely as a result of prefrontal cortex maturation (Diamond, 2002; Morton & Munakata, 2002). Further, children will need to monitor progress toward their goal and appraise whether a change in strategy is necessary.

Upon recognizing the failure of the current goal, children will need to evaluate other known strategies, the development of which are likely experience-dependent. For example, in our task, children have many potential strategies to choose from when it comes to selecting between the two games. They could choose randomly, switch back and forth between the two games, consistently choose the color they prefer, optimize performance via probability matching, or optimize performance by consistently choosing the easier game. The child may already know some of these strategies, from which they can select, but others may need to be discovered as currently known strategies prove ineffective (see Shrager & Siegler, 1998).

Finally, we think the ability to execute a known and selected strategy is more likely due to the maturation of prefrontal cortex during the early childhood years. This area has been implicated in the successful switching of attentional set by both children (in the DCCS; Morton, Bosma, & Ansari, 2009) and adults (in the Wisconsin Card Sorting Task; Buchsbaum, Greer, Chang, & Berman, 2005), and undergoes drastic development throughout early childhood as well as adolescence (Diamond, 2002).

Broader Implications of Current Findings

The reported research has broader theoretical implications for understanding cognitive development as well as practical implications related to learning and education. For example, our measure of metacognitive control can also be interpreted as a measure of top-down control, which holds relevance in many domains from attention to decision-making, and has been studied extensively in adults, children, and even primate species. Above, we described the dimensional change card sort (DCCS), which is a striking case of top-down control failure in early childhood, though there are plenty of other examples (e.g., delay of gratification, flanker task, Stroop task). In all of these cases, the experimenter must provide a strategy for the child to perform the task. In the current study, however, we were able to compare children's performance when it was necessary to formulate a strategy themselves to their performance when a strategy was provided. This allows us to broaden our understanding of top-down control to cases where this control is internally, rather than externally, initiated. Indeed, recent work has highlighted the distinction between proactive and reactive control processes, suggesting that metacognition may be key for engaging in proactive control in particular (Chevalier, Martis, Curran, & Munakata, 2015).

This work also has practical implications for teachers' use of feedback and strategy instruction in the classroom. These simple interventions improved children's metacognitive performance in the current task. However, it is important to investigate whether performance in the current task is transferrable to other tasks that measure metacognition. If performance in the current task transfers to performance in study time allocation tasks, this would have large implications for the training of metacognition. In particular, given that the present task is simpler and easier to administer than the study time allocation task, it could be used as a training platform for other more educationally relevant, yet difficult to administer tasks.

Questions for Future Research

Although several new findings stem from the reported research, some questions remain. Why was there no effect of feedback on children's metacognitive control? If feedback helped

children monitor the task difficulty, why did they not then use that knowledge to control behavior? There are several possibilities. One is that, although their monitoring improved, it did not improve *enough* to lead to better control. In other words, perhaps their representation of the task difficulty was not sufficiently robust. Alternatively, children may not have benefited because they had a hard time *tracking* and *remembering* the feedback they received. For example, they must track the occurrences of positive and negative feedback for easy and difficult trials across the task. Keeping this information in mind could create a rather large load on working memory, which may have made it difficult to access or use the information. Finally, it is possible that children are capable of using feedback to both monitor and control their behavior, but that they needed more evidence (e.g., more trials) to learn that some choices result in more negative feedback than others. Future work will assess the conditions under which children can benefit from performance feedback.

Conclusion

We demonstrated that, in contrast to adults, 5- and even 7-year-old children are not cognitive misers and do not minimize effort under typical circumstances. Perhaps more importantly, we found that metacognitive monitoring and control could be distinct processes, with each undergoing protracted development. At the same time, our findings suggest that the systems underlying these abilities are not *completely absent* in 5-year-olds. On the contrary, 5-year-olds demonstrated better monitoring when provided with feedback about their performance, and they were more likely to control their behavior when provided with a strategy. These findings provide novel evidence about the development of metacognition and its constituent processes.

References

- Bjorklund, D. F., Ornstein, P. A., & Haig, J. R. (1977). Developmental differences in organization and recall: Training in the use of organizational techniques. *Developmental Psychology, 13*(3), 175.
- Buchsbaum, B. R., Greer, S., Chang, W. L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the Wisconsin Card-Sorting task and component processes. *Human brain mapping, 25*(1), 35-45.
- Butler, R. (1990). The effects of mastery and competitive conditions on self-assessment at different ages. *Child Development, 61*(1), 201-210.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245-281.
- Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more: Executive function and symbolic representation in preschool children. *Psychological Science, 16*(8), 609-616.
- Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal of Cognitive Neuroscience, 27*(6), 1125-1136.
- Coughlin, C., Hembacher, E., Lyons, K. E., & Ghetti, S. (2014). Introspection on uncertainty and judicious help-seeking during the preschool years. *Developmental Science, 18*(6), 957-971.
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*, 2037-2078.
- Destan, N., Hembacher, E., Ghetti, S., & Roebers, C. M. (2014). Early metacognitive abilities: The interplay of monitoring and control processes in 5-to 7-year-old children. *Journal of Experimental Child Psychology, 126*, 213-228.

- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry. In D.T. Stuss & R.T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). New York: Oxford University Press.
- Dufresne, A., & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: Differential and sufficient aspects. *Journal of Experimental Child Psychology*, *47*, 274-296.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*, 906-911.
- Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.). (2009). *Handbook of metacognition in education*. Routledge.
- Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the “Number Sense”: The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*, 1457-1465.
- Hembacher, E. & Ghetti, S. (2014). Don't look at my answer: Subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science*, *25*, 1768-1776.
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679-709.
- Karpicke, J. D., Butler, A. C., & Roediger, H. L., III. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory (Hove, England)*, *17*, 471–479.
- Kendler, T. S. (1972). An ontogeny of meditational deficiency. *Child Development*, *43*, 1-17.
- Kool, W., & Botvinick, M. (2014). A labor/leisure tradeoff in cognitive control. *Journal of Experimental Psychology: General*, *143*, 131-141.
- Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, *139*, 665-682.

- Liberty, C., & Ornstein, P. A. (1973). Age differences in organization and recall: The effects of training in categorization. *Journal of Experimental Child Psychology, 15*, 169-186.
- Lockl, K., & Schneider, W. (2004). The effects of incentives and instructions on children's allocation of study time. *European Journal of Developmental Psychology, 1*, 153-169.
- Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development, 82*, 1778-1787.
- Lyons, K. E., & Ghetti, S. (2013). I don't want to pick! Introspection on uncertainty supports early strategic behavior. *Child Development, 84*, 726-736.
- Marcovitch, S., Boseovski, J. J., Knapp, R. J., & Kane, M. J. (2010). Goal neglect and working memory capacity in 4-to 6-year-old children. *Child Development, 81*, 1687-1695.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314-324.
- Metcalfe, J. E., & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. The MIT Press.
- Moely, B. E., Olson, F. A., Halwes, T. G., & Flavell, J. H. (1969). Production deficiency in young children's clustered recall. *Developmental Psychology, 1*, 26-34.
- Morton, J. B., Bosma, R., & Ansari, D. (2009). Age-related changes in brain activation associated with dimensional shifts of attention: an fMRI study. *Neuroimage, 46*, 249-256.
- Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology, 40*, 255-265.
<http://doi.org/10.1002/dev.10033>
- Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation, 26*, 125-173.
- Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19*, 90-138.
- Reder, L. M. (1996). *Implicit memory and metacognition*. Psychology Press.

- Roebbers, C. M. (2002). Confidence judgments in children's and adult's event recall and suggestibility. *Developmental Psychology, 38*, 1052.
- Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science, 1*, 291-297.
- Schneider, W. (2008). The development of metacognitive knowledge in children and adolescents: Major trends and implications for education. *Mind, Brain, and Education, 2*(3), 114-121.
- Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*, 405-410.
- Siegler, R. S., & Jenkins, E. A. (2014). *How children discover new strategies*. Psychology Press.
- Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling*. Hillsdale, NJ: Erlbaum
- Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 204-221.
- Vo, V.A., Li, R., Kornell, N., Pouget, A., & Cantlon, J.F. (2014). Young children bet on their numerical skills: Metacognition in the numerical domain. *Psychological Science, 25*, 1712-1721.
- Wiersema, J. R., van der Meere, J. J., & Roeyers, H. (2007). Developmental changes in error monitoring: An event-related potential study. *Neuropsychologia, 45*, 1649-1657.
- Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1594), 1310-1321.
- Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37-63.

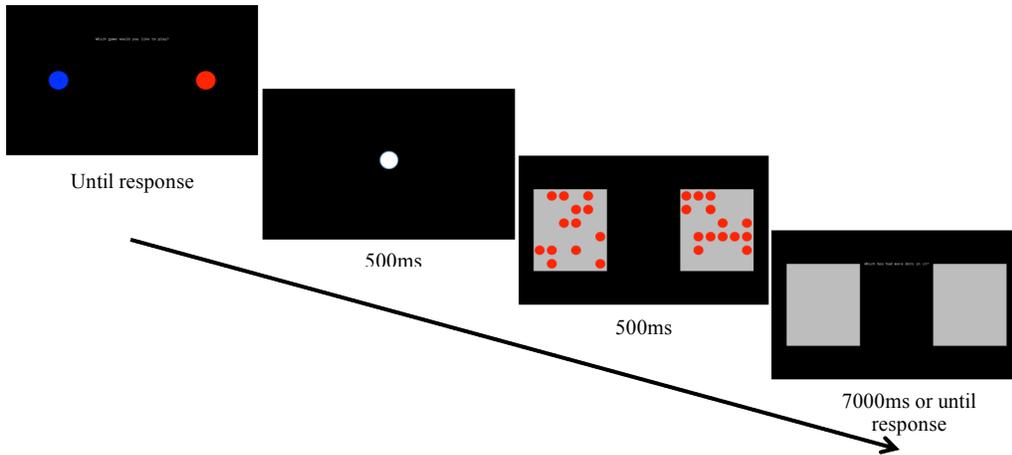


Figure 1. The task sequence including choice opportunity, fixation, test stimulus, and response screen.

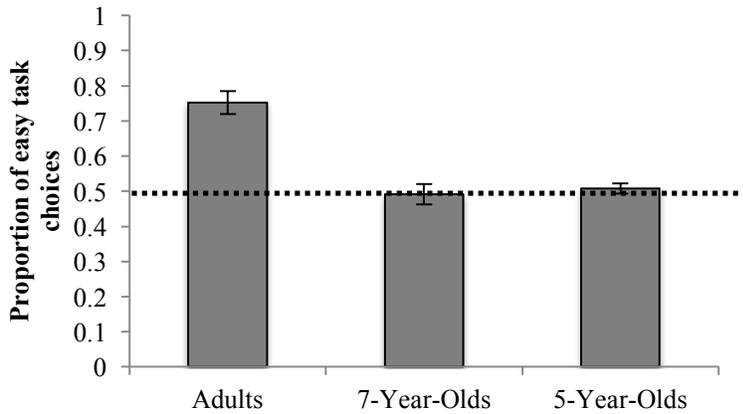


Figure 2. Proportion of easy task choices by age group.

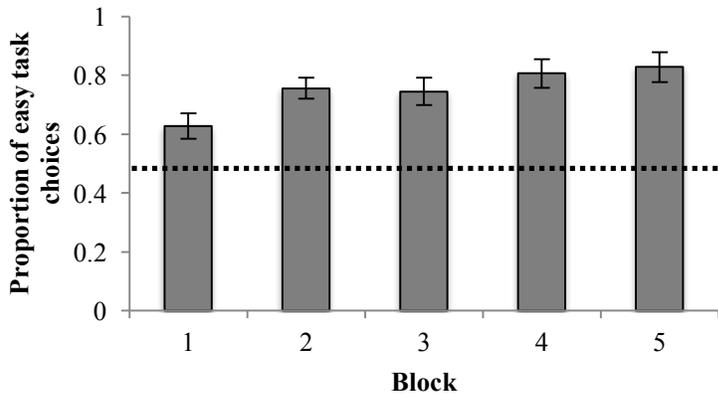


Figure 3. Proportion of adults' easy task choices by block.

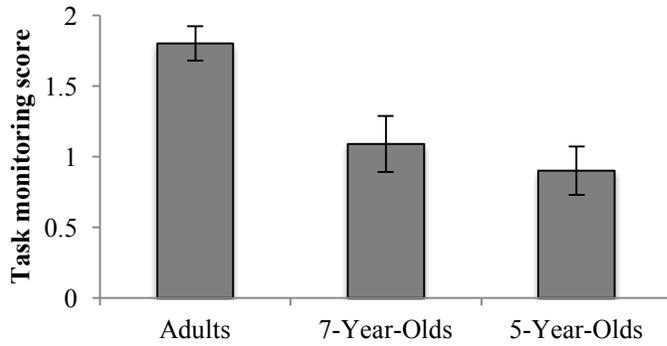


Figure 4. Composite task monitoring scores by age group.

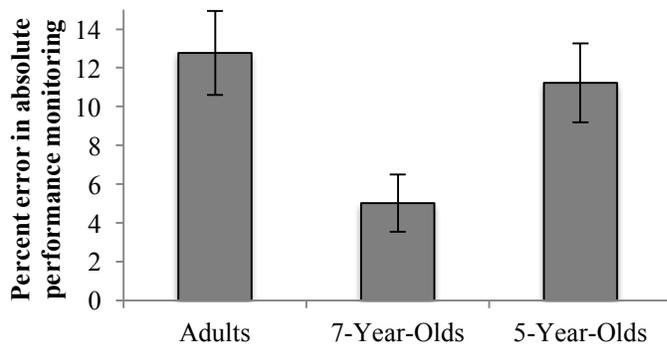


Figure 5. Absolute performance monitoring scores by age group (lower scores indicate greater performance estimation).

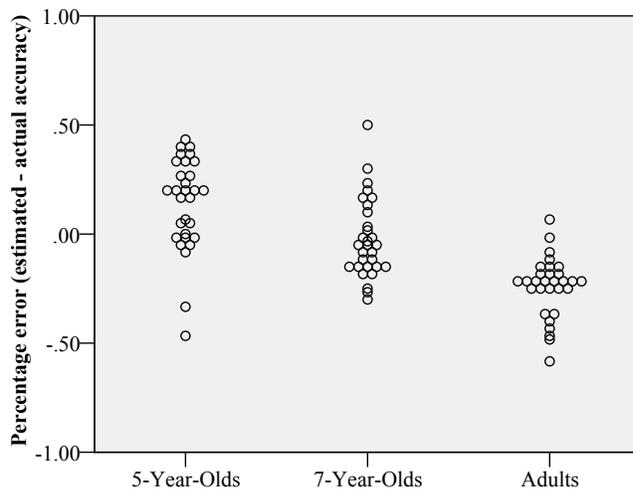


Figure 6. The direction of performance estimations for each age group. Scores above zero indicate overestimation.

Table 1: Summary of findings in Experiment 1.

	Adults	7-Year-Olds	5-Year-Olds
Discrimination accuracy			
Overall	0.92	0.82	0.74
Easy trials	1.00	0.99	0.89
Difficult trials	0.71	0.68	0.59
Discrimination RT (ms)			
Overall	900	825	1075
Easy trials	805	655	955
Difficult trials	1204	1032	1177
Control			
Easy task choices	0.75	0.49	0.51
Optimizers (out of 30)	70% ($N = 21$)	3% ($N = 1$)	3% ($N = 1$)
Task monitoring			
Composite score (out of 2)	1.88	1.09	0.90
Proficient monitors (out of 30)	77% ($N = 23$)	37% ($N = 11$)	17% ($N = 5$)
Performance monitoring			
Absolute	0.13	0.05	0.11
Relative (out of 30)	83% ($N = 25$)	50% ($N = 15$)	37% ($N = 11$)

Table 2: Summary of findings for 5-year-olds in Experiments 1 and 2.

	Baseline (Experiment 1)	Feedback Only	Instruction Only	Feedback + Instruction
Discrimination accuracy				
Overall	0.74	0.79	0.75	0.82
Easy trials	0.89	0.92	0.83	0.94
Difficult trials	0.59	0.65	0.62	0.58
Discrimination RT (ms)				
Overall	1075	1178	1684	1268
Easy trials	955	966	1566	1122
Difficult trials	1177	1252	1701	1486
Control				
Easy task choices	0.51	0.52	0.55	0.61
Optimizers (out of 30)	3% ($N=1$)	7% ($N=2$)	23% ($N=7$)	33% ($N=10$)
Task monitoring				
Composite score (out of 2)	0.90	1.19	1.30	1.69
Proficient monitors (out of 30)	17% ($N=5$)	43% ($N=13$)	43% ($N=13$)	67% ($N=20$)
Performance monitoring				
Absolute	0.11	0.12	0.13	0.10
Relative (out of 30)	37% ($N=11$)	37% ($N=11$)	37% ($N=11$)	40% ($N=12$)