American Psychological Association

Developmental Psychology

0012-1649/19/$12.00

2019, Vol. 55, No. 2, 315–328
http://dx.doi.org/10.1037/dev0000645

Components of Metacognition Can Function Independently Across Development

Allison P. O'Leary
Brevard College

Vladimir M. Sloutsky
The Ohio State University

It is often argued that metacognition includes 2 components: monitoring and control. However, it is unclear whether these components can operate independently, or whether they always operate as part of a hierarchy. The current study attempts to address this issue. In Experiment 1 ($N = 90$), age-related differences were assessed to examine the developmental trajectories of monitoring and control in 5- and 7-year-old children and adults. In Experiment 2 ($N = 90$) and Experiment 3 ($N = 90$), a scaffolding approach was taken with the same age groups to investigate correspondences in intervention-related changes in monitoring and control. Several dissociations between monitoring and control were found: In Experiment 2, strategy instruction affected metacognitive control, but not metacognitive monitoring, whereas in Experiment 3, performance feedback affected metacognitive monitoring, but not metacognitive control. These findings suggest that the monitoring and control components of metacognition can operate independently, challenging simple feed-forward models of metacognition.

Keywords: metacognition, strategy use, performance monitoring, cognitive development, learning

"Metacognition" refers to the ability to represent and access our own cognitive processes. This ability is crucial for understanding and optimizing how we learn, remember, and perform, allowing us to avoid strategies that have not worked for us in the past, and to continue using strategies that have. Consider a child who is studying for an upcoming science test. There are many different topics to learn and remember, and the child is more knowledgeable in some than in others. An efficient allocation of study time requires spending more time studying less familiar topics rather than distributing time equally among all the topics. To allocate time efficiently, it is necessary to (a) recognize that some topics are less familiar than others, (b) appreciate that less familiar topics will take more time and effort to learn, and (c) allocate study time according to the perceived difficulty of the topic. The abilities to evaluate one's own knowledge, proficiency, or performance (i.e., Points a and b above) have been referred to as *metacognitive monitoring*. The ability to adjust one's behavior in a goal-directed manner (i.e., Point c above) has been defined as *metacognitive control*. In this research we examine how these components of metacognition interact and whether these interactions change with development.

One form of metacognitive monitoring, *performance monitoring* is the ability to estimate one's success in a task and is often measured using judgments of confidence (i.e., how *certain* the participants are that they responded correctly) or accuracy estimations (e.g., how *many* correct responses were made). For both types of measures, young children vastly overestimate their performance (Roebers, 2002; Schneider, 1998). In contrast, adults' estimations are highly correlated with their actual performance, although they sometimes underestimate their performance (Devolder, Brigham, & Pressley, 1990; Koriat, Sheffer, & Ma'ayan, 2002; O'Leary & Sloutsky, 2017). Interestingly, 7-year-olds may be more accurate at evaluating their performance than both 5-year-olds and adults (O'Leary & Sloutsky, 2017), perhaps reflecting a transition from an overestimation bias in early childhood to an underestimation bias in adulthood.

Another component of metacognition is control—the ability to adjust one's behavior in accordance with their goals or task requirements. Such control often involves the formation and execution of a (presumably adaptive) strategy. *Strategy formation* requires the individual to (a) recognize that a current strategy is suboptimal and (b) select or devise a new strategy that is expected to be more effective (Shrager & Siegler, 1998). The ability to form strategies is expected to develop as individuals learn various strategies and ways of applying these strategies in different contexts (e.g., to allocate study time to lesser known topics, to space study time, or to prioritize an option that leads to higher reward). Notably, strategy formation may involve explicit, metacognitive knowledge of various strategies and their effectiveness (i.e., that selecting an easier task will save time and reduce errors).

Upon forming a strategy, the next step toward successful metacognitive control is to actually deploy the strategy. *Strategy execution* may involve the inhibition of a previously used, less effec-

This article was published Online First December 10, 2018.

Allison P. O'Leary, Department of Psychology, Brevard College; Vladimir M. Sloutsky, Department of Psychology, The Ohio State University.

This research was supported by National Institutes of Health Grants R01HD078545 and P01HD080679 and by Institute of Educational Sciences Grant R305A140214 to Vladimir M. Sloutsky.

Correspondence concerning this article should be addressed to Allison P. O'Leary, Department of Psychology, Brevard College, 1 Brevard College Drive, Brevard, NC 28712. E-mail: allison.p.oleary@gmail.com

tive strategy (e.g., to switch from indiscriminately selecting among task options to choosing a more beneficial option). Numerous studies have highlighted young children's difficulty executing a strategy even when the strategy is explicitly given (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004; Zelazo, Frye, & Rapus, 1996). As such, the ability to carry out a known strategy is thought to undergo dramatic development throughout early childhood (especially between 3 and 5 years of age) as well as adolescence (Lemaire & Brun, 2014; Zelazo, 2006), perhaps reflecting the development of prefrontal cortex (Morton, Bosma, & Ansari, 2009).

How Do Monitoring and Control Interact?

How do metacognitive monitoring and control interact? Does monitoring drive control? Does control drive monitoring? Do they operate independently? Does the relation between monitoring and control remain stable across development? The nature of the interaction has been a topic of debate. Below, we describe possible patterns of interaction between these metacognitive components: (a) monitoring drives (i.e., is a prerequisite of or *is necessary for*) control (MC); (b) control drives (is a prerequisite of or *is necessary for*) monitoring (CM); and (c) two variants of component independence (see Figure 1).

First, the monitoring drives control (MC) hypothesis posits a unidirectional, feed-forward relation between the components, such that proficient monitoring is a prerequisite for proficient control (Koriat, Ackerman, Adiv, Lockl, & Schneider, 2014; Koriat, Ma'ayan, & Nussinson, 2006; Nelson & Narens, 1990; Son, 2004). Under this interpretation, people need to detect (i.e., *monitor*) that some items are more difficult before formulating and executing the strategy to study the difficult items longer (i.e., *control* behavior). This approach suggests that (a) metacognitive

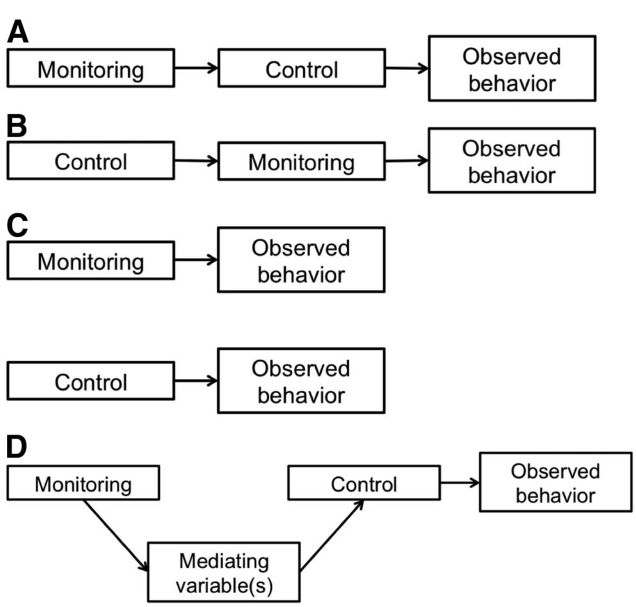

Figure 1. Possible relationships between monitoring and control. (A) MC model. (B) CM model. (C) Complete independence model. (D) Relative (mediated) independence model.

monitoring is necessary for metacognitive control and, therefore, (b) improvements in monitoring should *precede* improvements in control. Support for the MC model includes the finding that adults' judgments of learning (JOLs) during an initial study phase predicted which items they later selected for restudy (i.e., items rated as more poorly learned were restudied longer; Kornell & Metcalfe, 2006). As such, proponents of the MC model suggested that adults' appraisals of learning drove their selection of items for restudy. Koriat, Ackerman, Adiv, Lockl, and Schneider (2014) also found this correlation in older children when they were incentivized to maximize reward (e.g., by remembering items worth more points). In addition, both adults and older children allocate more study time to items that are judged to be difficult than to items judged to be easier (Dufresne & Kobasigawa, 1989; Dunlosky & Hertzog, 1998; Lockl & Schneider, 2004). Finally, some suggest this relation may hold across development, in that even young children seem able to act upon their metacognitive judgments (Destan, Hembacher, Ghetti, & Roebers, 2014).

A second possibility, the control drives monitoring (CM) hypothesis, suggests an opposite unidirectional effect, such that proficient control underlies proficient monitoring (Koriat & Ackerman, 2010; Koriat et al., 2014). Under this explanation, feedback from control operations (e.g., the amount of time or effort it takes to make a decision) is often the basis of metacognitive monitoring (e.g., evaluating how confident you are about that response). If this is the case, (a) successful control is necessary for successful monitoring, and (b) improvements in control should precede improvements in monitoring. This model is supported by studies in which participants' JOLs during test were *lower* for items that had been studied longer (Koriat et al., 2006). It was reasoned that, because participants had spent more time with those items (i.e., had found those items more difficult to commit to memory), they inferred they would be less likely to remember them in the future. Thus, monitoring is assumed to be based on the effort exerted from control processes. Evidence for this pattern has been observed in children (from first graders to eighth graders; Hoffmann-Biencourt, Lockl, Schneider, Ackerman, & Koriat, 2010; Koriat et al., 2014) as well as adults (Koriat et al., 2014). It should be noted, however, that there is evidence this relation is weaker (Hoffmann-Biencourt et al., 2010) or nonexistent (Koriat, Ackerman, Lockl, & Schneider, 2009) in younger children.

The third possibility is that monitoring and control can function independently (the component independence hypothesis). Under this construal, factors that influence monitoring may not influence control, and vice versa. For example, task variables (e.g., feedback) that improve children's monitoring performance may not improve their control performance. Previous work has shown that, in young children, metacognition is malleable, and improvements in monitoring can occur without corresponding improvements in control. As well, improvements in control may occur without improvements in monitoring (O'Leary & Sloutsky, 2017).

We consider two variants of the hypothesis: (a) *complete* independence and (b) *relative* independence. Complete independence presumes that changes in one component cannot lead to changes in the other component. However, this is unlikely to be the case. The classic work on the level of aspirations initially conducted by Hoppe (as cited in Irwin, 1944; Rotter, 1942) indicated that people are very adaptive in changing their own level of aspirations (measured by the level of task difficulty that they are willing to try) on the basis of their past performance on variants of the task. We, therefore, believe that relative independence is a more likely possibility. Under this construal, it is possible that the links between monitoring and control (in either the CM or MC direction) are mediated by other variables rather than being direct. For monitoring to influence control, for example, it may be necessary that the products of monitoring be converted into a strategy rule (one possible mediating factor). For example, knowing which of two tasks is easier may not influence one's task choice (i.e., control) unless it is accompanied by the strategy rule necessary for optimal performance: "select the easier task." In this case, the strategy rule could be generated either externally (e.g., given with task instructions) or internally (arrived at spontaneously by the individual), but is necessary for monitoring to affect control, and without it the components would appear independent (see Figure 1D for an example of the relative independence hypothesis).

Furthermore, the relation between monitoring and control can change throughout development, in that they may become more or less coupled with age and experience. For example, though the components may appear independent in early childhood, they may become more coordinated over time, as individuals gain experience using monitoring and control in tandem (e.g., using appraisals of performance to ask for help or to adjust study time). If this were the case, young children would demonstrate a pattern in line with the independence hypothesis, whereas adults (and, potentially, older children) would show a unidirectional pattern (in either the MC or CM direction).

The Current Study

The bulk of the research demonstrating interactions between monitoring and control has been descriptive, showing *relations* between monitoring and control ability (e.g., Koriat & Ackerman, 2010; Koriat et al., 2006, 2009, 2014). However, this approach has not allowed for strong causal explanations regarding the directions of these effects. In the current study, we designed an experimental approach to scaffold each component, and to observe the consequences of these manipulations on the other component. Our goal was to examine whether changes (i.e., improvements) in the monitoring component were *directly* related to changes in the control component, and vice versa. Both the MC and CM models predict that changes in the first component should precede changes in the second. If a manipulation improves only control (but not monitoring), this would provide evidence against the MC model. Similarly, if a manipulation improves monitoring, but not control, this would provide evidence against the CM model.

As such, the primary goal of the current study was to examine the coupling of monitoring and control across development. To achieve this goal, we used a task-choice paradigm (O'Leary & Sloutsky, 2017), in which on each trial participants decided which of two numerical discrimination "games" (i.e., blue dots or red dots) they wanted to perform. Unbeknownst to participants, each color corresponded to a level of discrimination difficulty (i.e., easy or difficult discriminations), with the color-difficulty contingency randomized between participants. The selection of a numerical estimation task minimizes developmental differences in performance (relative to a memory task, e.g., Halberda & Feigenson, 2008), thus reducing the possibility that metacognitive differences transpire due to differences in performance on the base-level task.

The ability to optimize performance by selecting the easier task was used as a measure of metacognitive control. This measure was selected based on previous work demonstrating that adults rely on control processes to avoid cognitive effort (e.g., the demand selection task used by Kool, McGuire, Rosen, & Botvinick, 2010). In addition, more recent work has used demand selection to assess metacognitive control in 6-year-old and 11-year-old children (Niebaum, Chevalier, Guild, & Munakata, 2018), showing a protracted developmental trajectory similar to other, classic metacognitive control measures (Dufresne & Kobasigawa, 1989).

The ability to evaluate one's own performance on the task was used as our measure of metacognitive monitoring. Because our base-level task (numerical discrimination) did not require learning, more classic measures like judgments of learning (JOLs) and ease of learning (EOLs) judgments could not be used. We also decided against using performance estimations (or confidence judgments) on *every* trial, to avoid prompting participants to monitor performance in a way they ordinarily would not. Repeatedly probing participants' monitoring could encourage interactions with control processes as an artifact of the task. Because we were specifically interested in spontaneous (rather than learned) links between monitoring and control, we measured monitoring only at the end of the task.

In Experiment 1, we examined the development of metacognitive monitoring and control (with the goal of replicating and extending prior findings) by measuring these abilities in 5-year-olds, 7-year-olds, and adults, in the absence of any scaffolding. We expected to see developmental improvements in both monitoring and control, across all three age groups.

Experiments 2 and 3 used a scaffolding approach to investigate component independence, by providing participants with strategy instruction (to improve control) and performance feedback (to improve monitoring), respectively. Providing scaffolding allowed us to directly assess whether the components are independent and whether this relation changes with development. Based on previous work (O'Leary & Sloutsky, 2017), we expected that young (and potentially older) children would show a pattern of data challenging both feed-forward models (both CM and MC). Other correlational studies with adults (Koriat et al., 2006, 2014), on the other hand, have been consistent with both the MC and CM accounts, so we entered the study with no preferred hypothesis for adults.

In addition, to investigate the deployment of metacognitive control (i.e., whether participants discovered and used a strategy rule), we focused on how participants' task choices changed over the course of the task. Across the experiments, we assessed whether participants adjusted their strategy (i.e., to select the easier task) by means of gradual or abrupt strategy change. We explain this approach in more detail in the results section of Experiment 1.

Experiment 1

Method

Participants. A sample of 5-year-olds (N = 30, 18 females, M = 5.35 years, SD = .25 years), 7-year-olds (N = 30, 12 females, M = 7.41 years, SD = .28 years), and undergraduate students from The Ohio State University (N = 30, 16 females, M = 19.98 years, SD = 1.93 years) participated in this experiment. In this and other experiments reported here, 5-year-olds were recruited through local daycares and preschools, 7-year-olds were recruited through local elementary schools in Columbus, Ohio, and adults were undergraduate students who received course credit for participation.

Materials and design. Stimuli were displayed on a Dell PC (for adults) or a Dell laptop connected to a Planar PXL2230MW 22-in. touch screen (for children) and controlled by OpenSesame presentation software (Mathôt, Schreij, & Theeuwes, 2012). Participants were presented with the choice task developed by O'Leary and Sloutsky (2017). On each trial, participants saw a pair of dots (i.e., one red and one blue) and were asked to select either the blue or the red game. They were then briefly presented with two sets of dots, both in the chosen color, and asked to identify which set had contained more dots. Unbeknownst to participants, the two colors were assigned to two levels of discrimination difficulty: easy and difficult (cf., Halberda & Feigenson, 2008). The assignment of color to difficulty was randomized across the participants but remained stable for each participant across the trials. Easy discriminations included dots presented in a 1:2 ratio and used the following values: four versus eight, five versus 10, six versus 12, seven versus 14, eight versus 16, nine versus 18, 10 versus 20, 11 versus 22, 12 versus 24, and 13 versus 26. Difficult discriminations included dots presented in a 9:10 ratio or smaller and included the following values: nine versus 10, 10 versus 11, 11 versus 12, 12 versus 13, and 13 versus 14.

Procedure. In this and other experiments reported here, participants were presented with procedures approved by The Ohio State University Institutional Review Board (Protocol #: 2004B042, *Comprehensive Protocol for Cognitive Development Research*). Adults performed the task in a lab on campus, whereas children performed the task in a quiet room in their preschool, daycare, or elementary school. At the beginning of the experiment, all participants were incentivized to perform as accurately as possible. Children (both 5- and 7-year-olds) were told that they would earn a point for each correct response, would lose a point for each incorrect response, and that their total number of points would determine how many stickers they would receive at the end of the game. Adults were told that they would earn 5 points for each correct response and would lose 5 points for each incorrect response. In addition, to encourage accurate performance, adult participants were told the number of points that would result in "above average," "average," or "below average" performance (in reality, the points were not tabulated). Each trial consisted of a choice opportunity, discrimination trial, fixation, and response screen (see Figure 2).

Measuring control. During the choice opportunity of each trial, participants were presented with a red and a blue dot, whose placement on the left or right of the screen were randomized. Participants were asked to select by either clicking (for adults) or using the touch screen (for children) which "game" (i.e., red or blue) they wanted to play. In so choosing, participants selected the corresponding difficulty level of the following discrimination trial. Importantly, participants were not explicitly told that the color mapped onto the difficulty level (nor that the games differed in difficulty at all) and had to learn this through experience with the games. Given that participants were incentivized to perform well, the optimal strategy was to consistently choose the easier game. Similar to previously reported work (see Kool et al., 2010), we used the proportion of easy task choices as a measure of metacognitive control.

Measuring discrimination performance. After participants selected the dot color of their choosing, a white fixation dot appeared for 500 ms, followed by the test stimulus. The test stimulus consisted of two gray boxes each containing a randomly positioned array of dots in the color the participant had just selected. These dot arrays were displayed according to the ratios described above (easy or difficult, depending on what color the participant chose), for 500 ms. After this, the dots disappeared, leaving only the empty gray boxes, and participants were asked to indicate by clicking (for adults) or touching (for children) which of the two boxes had contained more dots (see Figure 2). This response screen was displayed until the participant made a response, or until 7,000 ms had elapsed, after which the next trial began and participants were presented with another choice opportunity. Participants completed two practice trials before completing 30 test trials.

Measuring monitoring. Following the test trials, participants were asked three questions to assess their performance monitoring. Because we wanted to assess monitoring independently of assess-

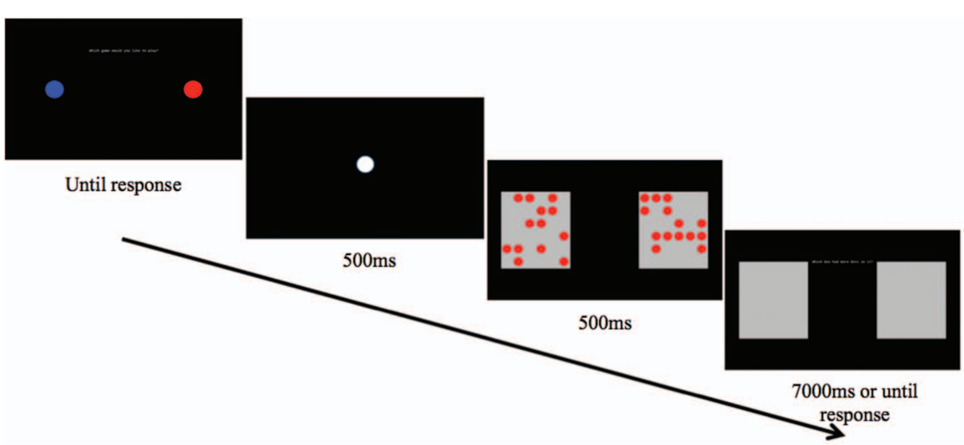

Figure 2. The task sequence including choice opportunity, fixation, test stimulus, and response screen. See the online article for the color version of this figure.

ing control, it was imperative to measure monitoring at the end of the task rather than throughout. This was done to avoid inducing performance reflections throughout the task that could have influenced control processes. To assess monitoring, participants were first asked to estimate how many trials they had correctly answered. Children were asked to choose from the following options: none of them, some of them, half of them, most of them, or all of them. Children indicated their answer by choosing a circle that was 0%, 25%, 50%, 75%, or 100% filled. Adults were asked to select the percentage of trials (from 0%, 25%, 50%, 75%, or 100%) that best corresponded to the amount they had correctly answered. By comparing each participant's estimate to their actual performance, we were able to assess their "absolute" performance monitoring, or how precisely they estimated their performance. Participants were also asked to make the same judgment for red trials only, and blue trials only (e.g., "How many of the [red/blue] ones did you get correct?"), the order of which was randomized. This allowed us to assess each participant's "relative" performance monitoring, or whether they recognized that their performance was higher on easy discrimination trials than on difficult trials. At the end of the experiment, all adult participants were told that their performance had been "excellent," and all children participants were awarded three stickers.

In Experiment 1, the main goal was to investigate the development of metacognitive monitoring and control in 5-year-olds, 7-year-olds, and adults in the absence of any scaffolding.

Results and Discussion

Discrimination accuracy. To verify that the two games were differentially difficult, we compared participants' performance in the easy and difficult games before proceeding with the main analyses. To assess effects of difficulty across the age groups, we conducted a 2 (Difficulty: Easy vs. Difficult) × 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) repeated-measures ANOVA on participants' discrimination performance. This analysis revealed a main effect of difficulty, in that participants were more accurate in the easy game (M = .92) than the difficult game (M = .65), $F(1, 83) = 123.05$, $p < .001$, $\eta^2 = .60$. There was also a main effect of age, $F(2, 83) = 22.55$, $p < .001$, $\eta^2 = .35$, in that both 7-year-olds (M = .84; $p < .001$) and adults (M = .85; $p < .001$) outperformed 5-year-olds (M = .66). Of particular interest was whether the level of difficulty affected the age groups similarly. Unexpectedly, there was a significant Difficulty × Age interaction, $F(2, 83) = 3.41$, $p < .05$, $\eta^2 = .08$, in that there was a smaller effect of task difficulty for 5-year-olds than the other two age groups (see Table 1). It should be noted, however, that this interaction was not found in Experiments 2 or 3, or in previous studies using this paradigm (O'Leary & Sloutsky, 2017). This may reflect anomalously poor performance in 5-year-olds' easy task performance in Experiment 1 (see Table 1). Despite the fact that this difference was *smaller* for 5-year-olds in this experiment, their performance was still significantly higher in the easy game than the difficult game, $t(29) = 4.80$, $p < .001$, $d = 1.10$.

Metacognitive control. Results are presented in Table 1 and Figures 3 and 4A. We used the proportion of participants' easy task choices as a measure of metacognitive control (see Figure 3A). As predicted, adults chose the easy game more often than would be expected by chance (M = 76%), $t(29) = 6.41$, $p < .001$,

Table 1
Summary of Results in Experiment 1

Experiment 1	5-year-olds	7-year-olds	Adults
Discrimination accuracy			
Overall	.66	.86	.93
Easy trials	.75	.99	1.00
Difficult trials	.57	.68	.69
Discrimination RT (ms)			
Overall	900	806	726
Easy trials	805	625	698
Difficult trials	1,204	937	973
Control			
Easy task choices	.51	.58	.75
Optimizers (out of 30)	7% (N = 2)	27% (N = 8)	66% (N = 20)
Performance monitoring			
Absolute (error)	.20	.05	.14
Relative (out of 30)	13% (N = 4)	30% (N = 9)	80% (N = 24)

$d = 2.38$. Seven-year-olds also systematically chose the easy game (M = 58%), $t(29) = 2.11$, $p < .05$, $d = .79$. Five-year-olds, however, did not (M = 51%), $p = .48$. A one-way ANOVA with age as a factor indicated that adults chose the easy game more often than both 5- and 7-year-olds, $F(2, 87) = 14.4$, $p < .001$, $\eta^2 = .25$ (see Table 1 and Figure 3A).

To assess individual differences in the task, we classified participants as "optimizers" if they systematically selected the easier task. To classify optimizers, we used a moving window of 12 trials (across the 30 trials in the task) to determine whether each participant chose the easier game on at least 11 of the 12 trials of *any* given window ($p = .052$, according to binomial probability). Proportions of optimizers across age groups are presented in Figure 3B. Twenty adults (66% of the sample) consistently chose the easy task, as did eight 7-year-olds (27% of the sample) and two 5-year-olds (7% of the sample; see Table 1). These proportions differed across the age groups, $\chi^2(2, N = 90) = 25.20$, $p < .001$, and post hoc comparisons revealed that more adults optimized than 7-year-olds, $\chi^2(1, N = 60) = 9.64$, $p < .005$, and 5-year-olds, $\chi^2(1, N = 60) = 23.25$, $p < .001$. In addition, more 7-year-olds optimized than 5-year-olds, $\chi^2(1, N = 60) = 4.32$, $p < .05$. Although 7-year-olds differed only numerically from 5-year-olds in their overall proportion of easy task choices, there were a significantly greater proportion of 7-year-olds than 5-year-olds who strategically chose the easy game. Taken together, these findings present evidence of metacognitive control development between 5- and 7-years-of-age, and even stronger evidence of development between 7-years-of-age and adulthood.

Because both adults and 7-year-olds chose the easy task more often than expected by chance (and both groups included substantial numbers of optimizers), we calculated backward learning curves to identify whether optimizers learned and adjusted their strategy gradually (indicating slow, associative learning) or abruptly (indicating rule discovery and application; Hayes, 1953). To calculate these curves, we first identified the trial at which each optimizer began to systematically choose the easier game. To do this, we again used a moving window of 12 trials to identify the earliest window at which each participant chose the easier game on at least 11 of these trials ($p = .052$, according to binomial probability). The first trial of this window was designated as Trial 0 (T_0). Identifying this trial allowed us to assess the rate of optimi-

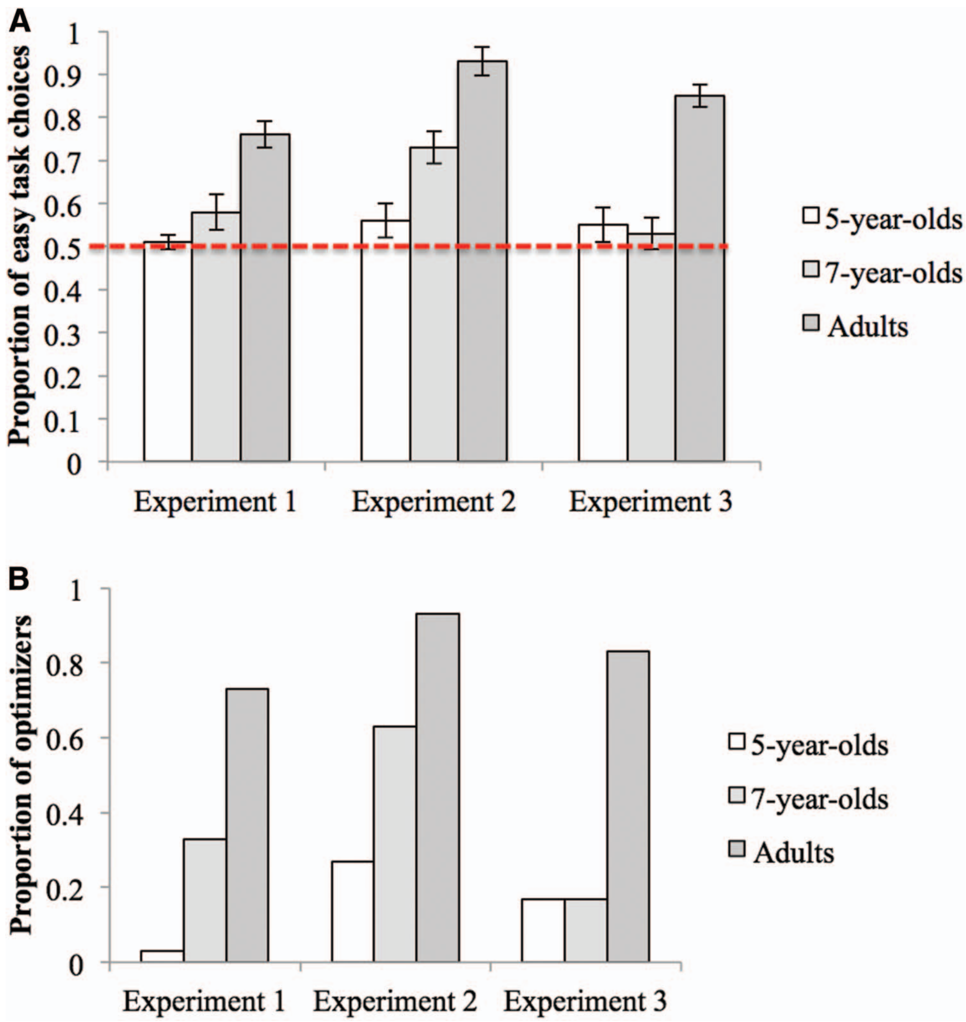

Figure 3. (A) Average proportions of easy task choices by age and experiment. (B) Proportions of optimizers (i.e., optimizing individuals) by age and experiment. Error bars represent SEMs. See the online article for the color version of this figure.

zation by aligning optimizers along the trial at which they began showing systematic metacognitive control. We then analyzed performance in blocks preceding and subsequent to T_0 (with five trials per block; see Figure 4A).

A shallow slope before $Block_0$ (i.e., the block containing T_0) coupled with a steep slope at T_0 (and reaching an asymptote before or during $Block_1$) would indicate that optimizers discovered the strategy rule and applied it in an all-or-nothing fashion (see Rehder & Hoffman, 2005, for related arguments). This pattern would suggest that once participants discovered which game would lead to optimal performance, they abruptly adjusted their strategy to consistently choose that game. This would also result in a shallow slope following optimization. In contrast, comparable slopes before, at, *and* after $Block_0$ would be indicative of associative (and perhaps more "implicit") learning rather than all-or-nothing rule discovery: As participants gradually learn associations between the color of the game and its corresponding outcome, they gradually adjust their responses.

For each optimizer, we calculated three slopes to indicate (a) learning prior to optimization (i.e., the slope between B_{-2} and B_{-1}); 9b) learning at T_0 (the slope between B_{-1} and B_0); and (c) learning following T_0 (the slope between B_0 and B_4). Due to the variability in the timing of T_0 across participants, some blocks were not populated for some optimizers (primarily the earliest blocks: B_{-1} and B_{-2}). Because these blocks represent the very beginning of the task for these individuals (i.e., there were no blocks before they showed choice optimization), it is reasonable to assume performance would have been at or around chance (.5). Thus, for each of these individuals, we calculated a performance estimate for these blocks of .5, jittered by a value between −.01 and .01. This allowed us to calculate the first two slopes mentioned above for every participant, while still maintaining some variability to allow for comparisons.

Adults who optimized ($N = 20$) in Experiment 1 demonstrated a steep slope at T_0 (B_{-1} to B_0; slope = .39, $t(19) = 10.35$, $p <$ .001; see Figure 4A). At the same time, there was no evidence of

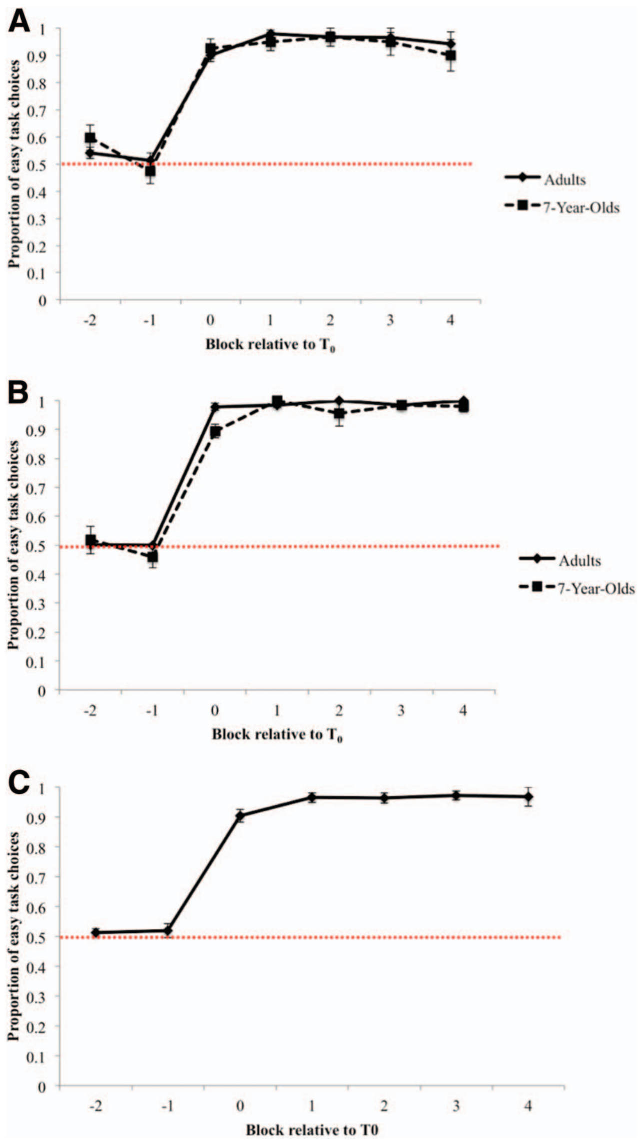

Figure 4. Backward learning curves in (A) Experiment 1 (for adults and 7-year-olds), (B) Experiment 2 (for adults and 7-year-olds), and C) Experiment 3 (for adults only). Error bars represent SEMs. Only adults are presented in Figure 4C because very few children optimized in Experiment 3. See the online article for the color version of this figure.

learning before or after optimization, in that the slope was not different from zero either before T_0 (B_{-2} to B_{-1}), slope = −.03, $t(19) = -.69$, $p = .50$, or after T_0 (B_0 to B_4), slope = .04, $t(13) = .90$, $p = .385$. These results indicate that adults adjusted their responses as soon as they discovered the relation between the color and task difficulty (consistent with use of a strategy rule), rather than adjusting their responses incrementally (which would be more consistent with associative learning).

The 7-year-olds who optimized ($N = 8$) displayed a very similar pattern of results, only showing substantial improvement at T_0 (B_{-1} to B_0; slope = .45, $t(7) = 6.18$, $p < .001$). This indicates that they, too, optimized behavior by applying an adaptive strategy rule. Slopes before and following T_0 did not differ from zero (both $ps > .18$). In sum, adult and 7-year-old optimizers showed a near identical pattern of performance, more consistent with a rapid strategy discovery than with gradual associative learning. Backward learning curves were not calculated for 5-year-olds because there were too few optimizers to calculate such curves.

Performance monitoring. Results are presented in Table 1 and Figures 5–6. To measure *absolute performance monitoring*, we compared each participant's performance estimates with their actual performance. To do this, we subtracted each participant's actual performance from their estimated performance, took the absolute value, and then adjusted these scores to account for our use of a discrete scale (see O'Leary & Sloutsky, 2017). Here, scores different from 0 indicated performance estimation error (see Figure 5A). Even adults significantly misestimated their performance ($M = .14$), $t(29) = 3.38$, $p < .005$, $d = 1.26$, demonstrating a strong tendency to underestimate (as demonstrated by their unadjusted estimated-actual performance scores; see Figure 6A). Five-year-olds showed the opposite pattern, in that they significantly overestimated their performance ($M = .20$; see Figures 4A and 5A), $t(29) = 7.50$, $p < .001$, $d = 2.79$. Seven-year-olds also misestimated performance ($M = .05$), $t(29) = 2.35$, $p < .05$, $d = .87$, but did so to a lesser extent than both 5-year-olds and adults, $F(2, 87) = 9.88$, $p < .001$, $\eta^2 = .18$. Although 7-year-olds' estimations were more accurate, they showed a trend toward underestimation, which was similar to adults (see Figure 6A).

We also assessed the proportion of participants who recognized that their performance was higher in the easy game than the difficult game (i.e., who successfully monitored their *relative* performance). As shown in Figure 5B, four 5-year-olds (13% of the sample), nine 7-year-olds (30% of the sample), and 24 adults (80% of the sample) rated their performance in the easy game as higher. These proportions were significantly different from one another, $\chi^2(2, N = 90) = 29.83$, $p < .001$. To assess the source(s) of the effect, we performed post hoc comparisons among the age groups. These analyses revealed that more adults monitored relative performance than either 5-year-olds, $\chi^2(1, N = 60) = 26.79$, $p < .001$, or 7-year-olds, $\chi^2(1, N = 60) = 15.15$, $p < .001$. Although the proportion of 7-year-olds exhibiting successful monitoring was numerically higher than that of 5-year-olds, these proportions did not differ significantly, $p = .12$.

Taken together, these findings indicate that performance monitoring develops between childhood and adulthood. Further, these data suggest that part of what changes throughout development is a bias in the *direction* of estimations, in that 5-year-olds show a tendency to overestimate performance, adults show a tendency to underestimate, and 7-year-olds represent a transitional group, already showing a slight tendency to underestimate.

Summary of Findings

Experiment 1 replicated and extended previous findings reported by O'Leary and Sloutsky (2017). Whereas adults spontaneously maximized performance and minimized effort by selecting an easier task, 7-year-olds did so to a lesser extent, and 5-year-olds did not at all. Seven-year-olds, however, more

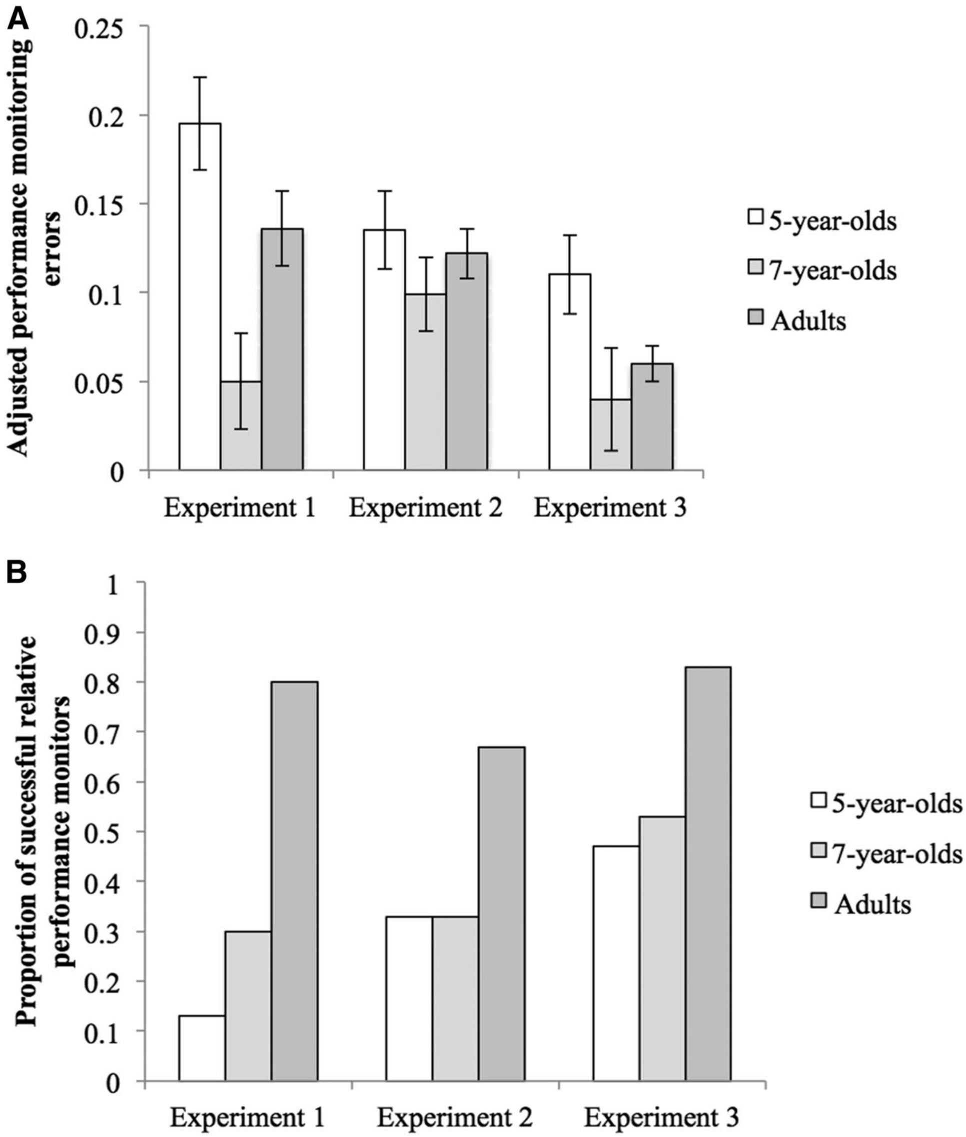

Figure 5. (A) Average adjusted performance monitoring errors (scores closer to 0 indicate more precise estimations) and (B) proportions of individuals exhibiting successful relative performance monitoring, both by age and experiment. Error bars represent SEMs.

accurately estimated their performance than both 5-year-olds and adults. Importantly, this difference reflected a transition between performance overestimation in early childhood, to performance underestimation in adulthood. In addition, relative performance monitoring showed some improvement between 5- and 7-years of age, as well as substantial improvement between childhood and adulthood. Finally, backward learning curves revealed that both 7-year-old and adult optimizers discovered the regularity (e.g., that the blue game was easier) and applied a strategy rule, rather than learning associations and gradually adjusting their behavior.

Experiment 1 highlights the processes of metacognitive monitoring and control when engaged spontaneously in the absence of any cue or instruction to perform optimally. In Experiment 2, a scaffolding approach was taken to directly investigate the relation between metacognitive monitoring and control. Specifically, we provided participants with a strategy (to choose the easier game) that would optimize their performance. Based on previous work (O'Leary & Sloutsky, 2017), we expected this scaffolding to lead to improvements in metacognitive control (but not monitoring) in 5-year-olds. However, effects of strategy instruction on monitoring and control in 7-year-olds and adults remain unknown. If improvements in control performance transpire in the absence of improvements in monitoring, this would provide evidence against the MC (monitoring is prerequisite of or it *drives* control) model of metacognition, which posits that improvements in monitoring should underlie improvements in control.

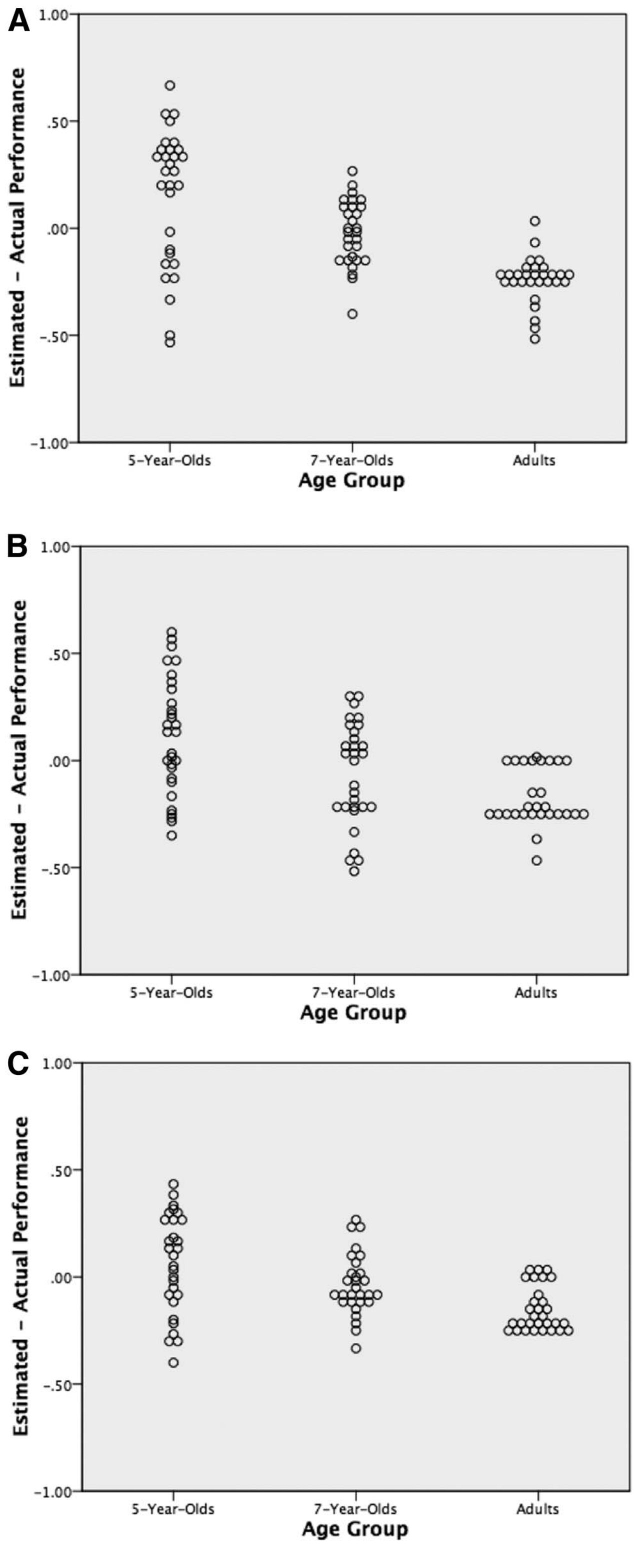

Figure 6. Distribution of individual performance estimation errors (estimated—actual) for 5-year-olds, 7-year-olds, and adults in (A) Experiment 1, (B) Experiment 2, and (C) Experiment 3. Each data point represents an individual. Points closer to 0 indicate more accurate performance estimations. Note that estimation errors are more tightly centered around 0 in Experiment 3 than in Experiment 1, indicating greater accuracy.

Experiment 2

Method

Participants. Thirty 5-year-olds ($N = 30$, nine females, $M = 5.25$ years, $SD = .22$ years), 30 7-year-olds ($N = 30$, 21 females, $M = 7.52$ years, $SD = .29$ years), and 30 undergraduate students ($N = 30$, 16 females, $M = 19.2$ years, $SD = 1.10$ years), participated in this experiment, none of whom participated in the previous experiment.

Materials, design, and procedure. The stimuli and procedure were similar to that of Experiment 1, with one exception: Participants were told at the beginning of the task that the two games differed in difficulty and were instructed, on each trial, to select the easier game (prior to each choice opportunity). Crucially, participants were not told which of the two games was easier.

Results and Discussion

Discrimination accuracy. To assess the effects of task difficulty across the age groups in Experiment 2, we again conducted a 2 (Difficulty: Easy vs. Difficult) × 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) repeated-measures ANOVA on participants' discrimination performance. As in Experiment 1, there was a main effect of difficulty, in that participants were more accurate in the easy game ($M = .92$) than the difficult game ($M = .59$), $F(1, 66) = 93.25$, $p < .001$, $\eta^2 = .59$. There was also a main effect of age, $F(2, 66) = 7.52$, $p < .005$, $\eta^2 = .19$, in that 5-year-olds ($M = .66$) were outperformed by both 7-year-olds ($M = .79$, $p < .005$) and adults ($M = .81$, $p < .005$). Most importantly, the Difficulty × Age interaction was not significant ($p = .10$), indicating that the level of difficulty affected performance similarly across the age groups.

Metacognitive control. Results are presented in Table 2 and Figures 3 and 4B. Similar to Experiment 1, adults ($M = 92\%$), $t(29) = 11.79$, $p < .001$, $d = 4.38$, and 7-year-olds ($M = 73\%$), $t(29) = 6.16$, $p < .001$, $d = 2.29$, chose the easier game more often than would be expected by chance, whereas 5-year-olds did not ($M = 57\%$; $p = .12$). A one-way ANOVA revealed an effect of age, $F(2, 87) = 21.97$, $p < .001$, $\eta^2 = .33$, in that adults

Table 2
Summary of Results in Experiment 2

Experiment 2	5-year-olds	7-year-olds	Adults
Discrimination accuracy			
Overall	.71	.89	.94
Easy trials	.80	.98	.97
Difficult trials	.55	.60	.63
Discrimination RT (ms)			
Overall	1,378	830	698
Easy trials	1,293	683	677
Difficult trials	1,516	1,125	1,160
Control			
Easy task choices	.57	.73	.92
Optimizers (out of 30)	20% ($N = 6$)	63% ($N = 19$)	93% ($N = 28$)
Performance monitoring			
Absolute (error)	.14	.10	.12
Relative (out of 30)	33% ($N = 10$)	33% ($N = 10$)	67% ($N = 20$)

significantly outperformed 7-year-olds ($p < .001$), who significantly outperformed 5-year-olds ($p < .005$), according to post hoc LSD comparisons (see Figure 3A).

As shown in Figure 3B, this pattern was also reflected in the proportion of *optimizers* (i.e., individuals who chose the easy game on at least 11 trials in a moving window of 12 trials) in each age group. A chi-square analysis revealed a significant effect of age, $\chi^2(2, N = 90) = 33.69$, $p < .001$. Post hoc comparisons showed that there were more adult optimizers ($M = 93\%$, 28 participants) than 7-year-old optimizers ($M = 63\%$, 19 participants), $\chi^2(1, N = 60) = 7.95$, $p < .01$, and more 7-year-olds who optimized than 5-year-olds ($M = 20\%$, 6 participants), $\chi^2(1, N = 60) = 11.59$, $p < .005$.

To examine how optimization was achieved, we again calculated backward learning curves for optimizers (see Figure 4B). Similar to Experiment 1, adult optimizers demonstrated a steep slope (slope $= .48$, $p < .001$) at T_0 (B_{-1} to B_0). All other slopes were not significantly different from zero. Seven-year-olds who optimized also showed a steep slope at T_0 (B_{-1} to B_0; slope $= .44$, $p < .001$), with no other slopes differing significantly from zero. Therefore, similar to Experiment 1, both adults and 7-year-olds exhibited abrupt rather than gradual strategy change, which is more consistent with discovering the regularity and abruptly changing the strategy than with gradual associative learning. Taken together, these findings suggest that the process of metacognitive control occurred similarly when participants were provided with a strategy (in Experiment 2) and when they had to formulate one spontaneously (in Experiment 1).

Performance monitoring. Results are presented in Table 2 and Figures 5–6. Participants in all age groups misestimated performance (with adjusted performance monitoring scores greater than 0; all $ps < .001$). There were no differences in participants' absolute performance monitoring across the three age groups in Experiment 2 ($p = .61$; see Figure 5A). In terms of the direction of their estimations, the pattern of findings was similar to that of Experiment 1, wherein adults underestimated performance, 5-year-olds overestimated performance, and 7-year-olds' estimations were more centered around 0.

Relative performance monitoring did differ across the age groups, $\chi^2(2, N = 90) = 9.00$, $p < .05$ (see Figure 5B), and post hoc comparisons revealed that more adults ($M = 67\%$) successfully monitored their relative performance than 7-year-olds ($M = 33\%$) and 5-year-olds ($M = 33\%$; both $ps < .05$). There was no difference between the proportions of 7-year-olds and 5-year-olds who successfully monitored their relative performance ($p = 1.00$).

Cross-experiment comparisons.

Metacognitive control. To directly assess effects of instruction on metacognitive control relative to Experiment 1, we conducted a 2 (Instruction: No Instruction vs. Instruction) × 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) ANOVA using participants' easy task choices from both Experiment 1 (where no instruction was provided) and Experiment 2 (with instruction). This analysis revealed a main effect of age, $F(2, 174) = 36.04$, $p < .001$, $\eta^2 = .29$. Post hoc LSD comparisons revealed that, overall, adults ($M = 92\%$) were more likely to select the easy game than both 7-year-olds ($M = 73\%$, $p < .001$) and 5-year-olds ($M = 56\%$, $p < .001$), and 7-year-olds were more likely to select the easy game than 5-year-olds ($p < .005$). More importantly, there was a main effect of instruction, $F(1, 174) = 17.64$, $p < .001$, $\eta^2 = .09$, in that participants in Experiment 2 ($M = 74\%$) outperformed those in Experiment 1 ($M = 62\%$). The interaction was not significant ($p = .25$), in that all age groups similarly benefited from instruction.

Performance monitoring. We conducted a 2 (Instruction: No Instruction vs. Instruction) × 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) ANOVA on participants' absolute performance monitoring scores in Experiments 1 and 2. This revealed only a significant main effect of age, $F(2, 174) = 6.88$, $p < .005$, $\eta^2 = .07$, in that 7-year-olds outperformed both 5-year-olds ($p < .001$) and adults ($p < .05$). There was no main effect of instruction ($p = .67$), nor was there a significant instruction by age interaction ($p = .09$), on participants' absolute performance monitoring scores.

We also assessed the proportion of participants who exhibited accurate relative performance monitoring (i.e., correctly detected that they had been more accurate in the easy game). Relative performance monitoring differed among the age groups in Experiments 1 and 2, $\chi^2(2, N = 180) = 35.18$, $p < .001$, in that more adults successfully monitored their relative performance than both 7-year-olds ($p < .001$), and 5-year-olds ($p < .001$). The proportion of 5- and 7-year-olds did not differ ($p = .31$). There was no overall effect of instruction on participants' relative performance monitoring, $p = .65$.

Summary of Findings

In Experiment 2, providing a strategy facilitated metacognitive control across the age groups, resulting in more systematic selection of an easier game. At the same time, there were no improvements in absolute or relative performance monitoring. In other words, the increases in easy task choices could not have stemmed from improved monitoring. This finding provides evidence against the MC model of metacognition.

In Experiment 3, we provided participants with performance feedback to test the CM hypothesis. Previous work has shown that 5-year-olds' monitoring benefited from feedback (O'Leary & Sloutsky, 2017), likely by providing an external signal of performance in the task. As suggested by the findings of Experiment 1, 5-year-olds have difficulty spontaneously estimating performance, which may lead them to rely on explicit feedback to gauge their performance. Improvements in control that correspond to improvements in monitoring would provide evidence for interdependence via the CM model. However, improvements in monitoring in the absence of improvements in control would further support the hypothesis that the two processes can operate independently.

In addition to providing an external signal of performance, feedback perhaps makes it more likely that participants will rely on associative learning to select the easier game. For example, instead of forming a strategy rule (e.g., "I should choose the blue game because it is easier") they may form a more implicit representation of response-outcome contingencies, based on the type of feedback (i.e., positive or negative) received following the selection of each game. As discussed before, this type of learning should result in participants' learning curves showing a more gradual increase. Application of a strategy or rule, on the other hand, should result in the profile observed in adults and 7-year-olds in Experiments 1 and 2.

Experiment 3

Method

Participants. 5-year-olds ($N = 30$; 11 females, $M = 5.41$ years, $SD = .28$ years), 7-year-olds ($N = 30$; 14 females, $M = 7.50$ years, $SD = .32$ years), and undergraduate students ($N = 30$; 16 females, $M = 19.42$ years, $SD = 1.55$ years) participated in this experiment, none of whom participated in any of the previous experiments.

Materials, design, and procedure. The stimuli and procedure were similar to those in Experiment 1, with one exception. In this experiment, participants received feedback about their performance after each discrimination response. If they answered correctly, they saw a smiley face and heard a high tone for 500 ms. If they answered incorrectly, they saw a sad face and heard a low tone for 500 ms.

Results and Discussion

Discrimination accuracy. We performed a 2 (Difficulty: Easy vs. Difficult) × 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) repeated-measures ANOVA on participants' discrimination accuracy to assess the effect of difficulty across the age groups in Experiment 3. As before, there was a main effect of difficulty, $F(1, 81) = 190.25$, $p < .001$, $\eta^2 = .70$, in that participants were more accurate in the easy game ($M = .95$) than the difficult game ($M = .67$). There was also a main effect of age, $F(2, 81) = 7.32$, $p < .005$, $\eta^2 = .15$, in that 5-year-olds ($M = .74$) performed more poorly than both 7-year-olds ($M = .82$, $p < .05$) and adults ($M = .87$, $p < .001$) overall. Finally, and most importantly, the interaction between difficulty and age was not significant ($p = .23$), indicating that the level of difficulty affected discrimination performance similarly across the age groups.

Metacognitive control. Results are presented in Table 3 and Figures 3 and 4C. As shown in Figure 3A, only adults in Experiment 3 chose the easy task more than expected by chance, 85%, $t(59) = 12.43$, $p < .001$, $d = 3.24$, whereas 7-year-olds ($M = 53\%$) and 5-year-olds ($M = 55\%$) did not (both $ps > .12$). A one-way ANOVA revealed an effect of age on the proportion of easy task choices, $F(2, 87) = 35.16$, $p < .001$, $\eta^2 = .45$. Post hoc LSD comparisons indicated this difference was due to the fact that adults outperformed both 5-year-olds ($p < .001$) and 7-year-olds ($p < .001$).

This pattern was also reflected in the proportions of participants who were classified as optimizers (see Figure 3B). A chi-square analysis revealed an effect of age, $\chi^2(2, N = 90) = 35.53$, $p < .001$, and post hoc comparisons showed that this was due to adults outperforming both 7-year-olds ($p < .001$) and 5-year-olds ($p < .001$).

As in previous experiments, we calculated backward learning curves for the adults who systematically chose the easier game (curves were not calculated for 7-year-olds because there were too few optimizers; see Figure 4C). As in previous experiments, the slope at T_0 (B_{-1} to B_0) was steep (slope = .39, $p > .001$, above 0), suggesting abrupt strategy rule discovery. However, in contrast to Experiments 1 and 2, there was a small, yet nonzero slope after T_0 (B_0 to B_4; slope = .07, above 0, $p < .05$). Though this slope was nonzero, it was substantially smaller than the average slope at T_0, $t(18) = 6.54$, $p < .001$, $d = 1.54$. Overall, the profile of the backward learning curve was very similar to those observed in Experiments 1 and 2. Notably, participants who did not optimize showed no improvement in their easy task choices across the testing session, and this pattern did not differ as a function of age or experiment (all $ps < .09$; see Figure 7). This further supports our claim that participants did not gradually learn and use the contingency between the task color and task difficulty.

Performance monitoring. Results are presented in Table 3 and Figures 5–6. As in previous experiments, participants of all age groups misestimated performance, with performance monitoring errors differing from 0; all $ps < .01$ (see Figure 5A). Participants' absolute performance monitoring differed as a function of age, according to a one-way ANOVA, $F(2, 87) = 5.34$, $p < .01$, $\eta^2 = .11$. Post hoc LSD comparisons revealed that both adults ($M = .06$) and 7-year-olds ($M = .04$) more accurately estimated their performance than 5-year-olds ($M = .11$; both $ps < .05$). Similar to previous experiments, adults underestimated performance, 5-year-olds overestimated, and 7-year-olds estimations were more centered around 0 (Figure 6C). It should be noted, however, that participants' estimates were somewhat more compressed around 0 in this experiment than in the previous experiments.

Results of relative performance monitoring are presented in Figure 5B. A chi-square analysis also revealed an effect of age on the proportion of participants who accurately monitored their relative performance, $\chi^2(2, N = 90) = 9.63$, $p < .01$. Post hoc comparisons revealed that adults outperformed 7-year-olds ($p < .05$) and 5-year-olds ($p < .005$).

Cross-experiment comparisons.

Metacognitive control. As in Experiment 2, we conducted a 2 (Feedback: No Feedback vs. Feedback) × 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) ANOVA on the proportion of easy task choices in Experiments 1 and 3. This analysis revealed only a main effect of age, $F(2, 174) = 44.86$, $p < .001$, $\eta^2 = .34$, in that adults outperformed both 5-year-olds ($p < .001$) and 7-year-olds ($p < .001$; see Figure 3A). There was no main effect of feedback on metacognitive control ($p = .31$; $\eta^2 = .01$), nor did the interaction between age and feedback reach significance ($p = .09$, $\eta^2 = .03$).

Performance monitoring. We conducted a 2 (Feedback: No Feedback vs. Feedback) × 3 (Age: 5-year-olds vs. 7-year-olds vs.

Table 3
Summary of Results in Experiment 3

Experiment 3	5-year-olds	7-year-olds	Adults
Discrimination accuracy			
Overall	.76	.83	.96
Easy trials	.91	.95	1.00
Difficult trials	.58	.70	.74
Discrimination RT (ms)			
Overall	1,230	1,092	653
Easy trials	1,050	936	612
Difficult trials	1,322	1,166	932
Control			
Easy task choices	.55	.53	.85
Optimizers (out of 30)	13% ($N = 4$)	13% ($N = 4$)	77% ($N = 23$)
Performance monitoring			
Absolute (error)	.11	.04	.06
Relative (out of 30)	47% ($N = 14$)	53% ($N = 16$)	83% ($N = 25$)

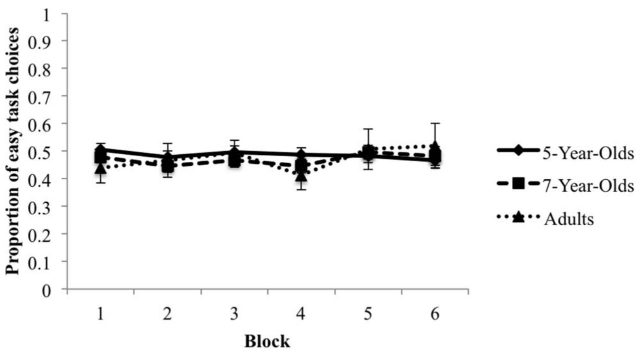

Figure 7. Proportions of easy task choices made by nonoptimizers across Experiments 1, 2, and 3. Performance was broken into blocks of five trials.

Adults) ANOVA on participants' absolute performance monitoring scores in Experiments 1 and 3. There was a main effect of feedback, $F(1, 174) = 12.85$, $p < .001$, $\eta^2 = .07$, in that performance estimations were more accurate in Experiment 3 ($M = .07$) than in Experiment 1 ($M = .13$). There was also a main effect of age, $F(2, 174) = 14.86$, $p < .001$, $\eta^2 = .15$, in that 7-year-olds performance estimates were more accurate than both 5-year-olds' ($p < .001$) and adults' ($p < .01$). Adults' estimations were more accurate than 5-year-olds' ($p < .01$). The feedback by age interaction was not significant ($p = .12$), suggesting that feedback influenced metacognitive monitoring similarly across the three age groups.

We also assessed the effects of feedback on the proportion of participants who exhibited successful relative performance monitoring. Most importantly, there was a significant effect of feedback, $\chi^2(1, N = 180) = 7.20$, $p < .01$. There was also a significant overall effect of age, $\chi^2(2, N = 180) = 35.26$, $p < .001$, in that more adults successfully monitored their relative performance than 7-year-olds ($p < .001$) and 5-year-olds ($p < .001$). The proportion of successful relative performance monitors did not differ between 5 and 7 years of age ($p = .18$).

Summary of Findings

In Experiment 2, strategy instruction improved participants' metacognitive control performance, in the absence of improvements in metacognitive monitoring, providing evidence against the MC model of metacognition. In contrast, we found the opposite tendency in Experiment 3: Participants' performance monitoring improved with performance feedback, whereas control performance was unaffected. This provides evidence against the CM model, in that increased precision of performance monitoring did not rely on changes in control performance.

How did adults learn the link between the color and task difficulty in Experiment 3? As suggested at the end of Experiment 2, it was possible that participants would learn which task was easier through slowly accumulated associations between the task type (i.e., red or blue) and the performance feedback (i.e., correct or incorrect). On the contrary, backward learning curves indicated that feedback did not encourage associative learning; instead, as in the previous experiments, adults appeared to abruptly discover and apply a strategy rule rather than learn it gradually.

General Discussion

In three experiments, we investigated the developmental trajectory of the interaction between metacognitive monitoring and control. In Experiment 1, we measured the developmental time course of monitoring and control. In subsequent experiments we considered possible links between monitoring and control. Specifically, we considered three possibilities: (a) monitoring driving control (MC); (b) control driving monitoring (CM); and (c) component independence. Though previous work provided correlational evidence for both a MC and a CM model of metacognition, we directly manipulated the monitoring and control components in Experiments 2 and 3, to test each component's direct impact upon the other.

In Experiment 2, we tested the possibility that metacognition follows a MC model, which suggests that control processes are guided by monitoring processes. Under this model, improvements in control should occur only if there are improvements in monitoring. To test this possibility, we provided 5-year-olds, 7-year-olds, and adults with a strategy to employ in the game (i.e., reducing the need to formulate a strategy spontaneously), which has been shown to improve metacognitive control in young children (O'Leary & Sloutsky, 2017). We found evidence for improvements in control without improvements in monitoring across the three age groups, providing evidence against the MC model.

In Experiment 3, we used performance feedback to scaffold participants' performance monitoring and to test the CM hypothesis. Specifically, we assessed whether improvements in monitoring were driven by improvements in control. In this experiment, we observed improvements in both measures of performance monitoring, in the absence of control improvements. This provides evidence against the CM hypothesis, and further suggests that monitoring and control can operate independently across all three age groups.

Overall, the current data replicate and substantially extend previous findings supporting the independence model of metacognition (O'Leary & Sloutsky, 2017). The reported results suggest that monitoring and control can operate independently, and that the two processes do not become more coupled with experience and development. Below, we discuss these findings and their implications for the relation between the monitoring and control processes of metacognition.

Can Monitoring and Control Function Independently?

Previous work has provided correlational evidence that monitoring and control are interactive. Some evidence suggests that monitoring guides control (i.e., the MC hypothesis). For example, low initial judgments of learning predicted longer restudy times, meaning that people tend to allocate more study time to those items about which they initially had low confidence (Kornell & Metcalfe, 2006). Other work suggests the opposite contingency—that feedback from control processes guides metacognitive monitoring (i.e., the CM hypothesis).

In the reported experiments, we found a number of challenges to the idea of direct, feed-forward links between monitoring and control. In Experiment 2, strategy instruction improved control, with no corresponding changes in monitoring. In Experiment 3, feedback influenced monitoring, whereas control was unaffected. Importantly, these dissociations persisted across the three tested age groups, suggesting that coupling does not increase with development, and that monitoring and control can operate independently across development.

Our conclusion that the components of metacognition can operate independently stems from evidence that change in one component neither resulted in, nor required, change in the other component. This conclusion, however, raises important theoretical questions. If monitoring and control operate independently, how does this independence manifest itself? From the very beginning, we considered two possibilities: (a) complete independence; and (b) relative independence that stems from indirect (i.e., mediated) links between the components of metacognition. We believe that the complete independence is rather unlikely: As discussed above, the classic work on the level of aspirations initially conducted by Hoppe (see Irwin, 1944; Rotter, 1942) indicated that people are adaptive in changing their own level of aspirations on the basis of their past performance on variants of the task. Thus, we believe that *relative* independence (i.e., mediated links) is more likely, in that the interaction between monitoring and control is mediated by other variables rather than being direct. For example, successfully monitoring one's performance may not influence control unless the individual is able to convert that knowledge into a strategy rule. Indeed, by assessing the backward learning curves of 7-year-olds and adults, we found those who successfully controlled their behavior did so by updating their strategy in an abrupt, all-or-nothing fashion. In other words, it appears that participants needed to convert their internal signal of difficulty (e.g., performance estimations) into a strategy (e.g., to select the easier task), to effectively control behavior. However, it is possible that under some conditions this conversion may fail (e.g., due to a weak internal signal or lack of metacognitive knowledge of which strategy may work), in which case changes in monitoring would not result in an updated strategy. Under this construal, it also possible that a strong external signal to change strategy may result in a strategy change, without changes in monitoring.

Though we suggest strategy rule formation as a possible mediating factor, the current study does not directly test the mediated link hypothesis and so additional research is needed. At the same time, the current work strongly challenges the idea that monitoring and control are completely dependent, such that change in one component is necessary to achieve change in another component.

Limitations of the Current Study

One potential limitation of the current study is that we used a simple perceptual task as our base-level task, whereas many previous studies have used learning tasks. However, we believe that monitoring memory performance would recruit similar processes as those required to monitor discrimination performance. Indeed, the developmental trajectory for uncertainty monitoring is similar across perceptual, lexical, and memory tasks (Hembacher & Ghetti, 2014; Lyons & Ghetti, 2011). Additionally, young children have shown similar patterns of metacognitive control in tasks involving both help seeking (Coughlin, Hembacher, Lyons, & Ghetti, 2015) and response withholding (Hembacher & Ghetti, 2014). These findings suggest that the type of base-level task should have little effect on the metacognitive processes involved; however, the generality of the current findings needs to be tested with a range of monitoring and control tasks.

In addition, numerical discrimination performance differed somewhat across the age groups. Though 7-year-olds and adults' performance was comparable, 5-year-olds were less accurate. As such, it is possible that these base-level performance variations could have influenced metacognitive performance. However, the facts that (a) performance differences between the easy and the difficult tasks were comparable among the age groups, and (b) participants of all three groups responded equivalently to experimental manipulations, mitigate against such concerns. At the same time, as suggested by Experiments 1–3, equivalent task difficulty does not guarantee equivalent monitoring—children's relative performance monitoring was consistently below that of adults. This finding points to important development differences in monitoring that are not driven by differences in the base-level task difficulty. The sources of these difficulties in monitoring have to be further examined in future research.

Another potential limitation is that we measured metacognitive monitoring only at the end of the task. This was done to ensure that frequent prompts to reflect on the task did not bias participants' metacognitive control. However, one could argue that measuring monitoring in this way may result in insufficient power (relative to JOLs or confidence judgments, which are typically measured on each trial). Therefore, more quantitatively precise measures of monitoring, which do not interfere with measures of control, need to be developed in future research.

Conclusion

This research presents novel findings suggesting that metacognition undergoes protracted development. Most importantly, we presented evidence of dissociations between monitoring and control across development, providing evidence against the CM (control → monitoring) and MC (monitoring → control) models of metacognition. These findings suggest that monitoring and control can function relatively independently from early childhood, and that the two components do not seem to become more directly coupled with experience and development.

References

Coughlin, C., Hembacher, E., Lyons, K. E., & Ghetti, S. (2015). Introspection on uncertainty and judicious help-seeking during the preschool years. *Developmental Science, 18,* 957–971. http://dx.doi.org/10.1111/desc.12271

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. http://dx.doi.org/10.1016/j.jecp.2014.04.001

Devolder, P. A., Brigham, M. C., & Pressley, M. (1990). Memory performance awareness in younger and older adults. *Psychology and Aging, 5,* 291–303. http://dx.doi.org/10.1037/0882-7974.5.2.291

Dufresne, A., & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: Differential and sufficient aspects. *Journal of Experimental Child Psychology, 47,* 274–296. http://dx.doi.org/10.1016/0022-0965(89)90033-7

Dunlosky, J., & Hertzog, C. (1998). Aging and deficits in associative memory: What is the role of strategy production? *Psychology and Aging, 13,* 597–607. http://dx.doi.org/10.1037/0882-7974.13.4.597

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. http://dx.doi.org/10.1037/a0012682

Hayes, K. J. (1953). The backward curve: A method for the study of learning. *Psychological Review, 60,* 269–275. http://dx.doi.org/10.1037/h0056308

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. http://dx.doi.org/10.1177/0956797614542273

Hoffmann-Biencourt, A., Lockl, K., Schneider, W., Ackerman, R., & Koriat, A. (2010). Self-paced study time as a cue for recall predictions across school age. *British Journal of Developmental Psychology, 28,* 767–784. http://dx.doi.org/10.1348/026151009X479042

Irwin, F. W. (1944). The realism of expectations. *Psychological Review, 51,* 120–126. http://dx.doi.org/10.1037/h0061743

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. http://dx.doi.org/10.1037/a0020198

Koriat, A., & Ackerman, R. (2010). Metacognition and mindreading: Judgments of learning for Self and Other during self-paced study. *Consciousness and Cognition, 19,* 251–264. http://dx.doi.org/10.1016/j.concog.2009.12.010

Koriat, A., Ackerman, R., Adiv, S., Lockl, K., & Schneider, W. (2014). The effects of goal-driven and data-driven regulation on metacognitive monitoring during learning: A developmental perspective. *Journal of Experimental Psychology: General, 143,* 386–403. http://dx.doi.org/10.1037/a0031768

Koriat, A., Ackerman, R., Lockl, K., & Schneider, W. (2009). The memorizing effort heuristic in judgments of learning: A developmental perspective. *Journal of Experimental Child Psychology, 102,* 265–279. http://dx.doi.org/10.1016/j.jecp.2008.10.005

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135,* 36–69. http://dx.doi.org/10.1037/0096-3445.135.1.36

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General, 131,* 147–162. http://dx.doi.org/10.1037/0096-3445.131.2.147

Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 609–622. http://dx.doi.org/10.1037/0278-7393.32.3.609

Lemaire, P., & Brun, F. (2014). Effects of strategy sequences and response-stimulus intervals on children's strategy selection and strategy execution: A study in computational estimation. *Psychological Research, 78,* 506–519. http://dx.doi.org/10.1007/s00426-013-0501-0

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. http://dx.doi.org/10.1080/17405620444000085

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development, 82,* 1778–1787. http://dx.doi.org/10.1111/j.1467-8624.2011.01649.x

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44,* 314–324. http://dx.doi.org/10.3758/s13428-011-0168-7

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. http://dx.doi.org/10.1016/j.neuroimage.2009.01.037

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development, 75,* 1850–1870. http://dx.doi.org/10.1111/j.1467-8624.2004.00821.x

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation, 26,* 125–173. http://dx.doi.org/10.1016/S0079-7421(08)60053-5

Niebaum, J. C., Chevalier, N., Guild, R. M., & Munakata, Y. (2018). Adaptive control and the avoidance of cognitive control demands across development. *Neuropsychologia*. Advance online publication. http://dx.doi.org/10.1016/j.neuropsychologia.2018.04.029

O'Leary, A. P., & Sloutsky, V. M. (2017). Carving metacognition at its joints: Protracted development of component processes. *Child Development, 88,* 1015–1032. http://dx.doi.org/10.1111/cdev.12644

Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology, 51,* 1–41. http://dx.doi.org/10.1016/j.cogpsych.2004.11.001

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75,* 1387–1401. http://dx.doi.org/10.1111/j.1467-8624.2004.00747.x

Roebers, C. M. (2002). Confidence judgments in children's and adults' event recall and suggestibility. *Developmental Psychology, 38,* 1052–1067. http://dx.doi.org/10.1037/0012-1649.38.6.1052

Rotter, J. B. (1942). Level of aspiration as a method of studying personality. I. A critical review of methodology. *Psychological Review, 49,* 463–474. http://dx.doi.org/10.1037/h0057506

Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science, 1,* 291–297. http://dx.doi.org/10.1111/1467-7687.00044

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9,* 405–410. http://dx.doi.org/10.1111/1467-9280.00076

Son, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 601–604. http://dx.doi.org/10.1037/0278-7393.30.3.601

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1,* 297–301. http://dx.doi.org/10.1038/nprot.2006.46

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11,* 37–63. http://dx.doi.org/10.1016/S0885-2014(96)90027-1

Received April 24, 2017
Revision received September 8, 2018
Accepted September 12, 2018 ■